

```

1 CTCTTACTCT TCAGCCTGAT GTCAAAAGCA AAAGTTCAGA AGTTCCTCAT
51 CAATAAGGAG TCCTTGTGAG CAGGTGAAGC TCATCTAACT AGGCATTTCT
101 ATGATGTGGC TGCTTTTAAC AACAACTTGT TTGATCTGTG GAACTTTAAA
151 TGCTGGTGGA TTCCTTGATT TGGAAAATGA AGTGAATCCT GAGGTGTGGA
201 TGAATACTAG TGAAATCATC ATCTACAATG GCTACCCAG TGAAGAGTAT
251 GAAGTCACCA CTGAAGATGG GTATATACTC CTTGTCAACA GAATTCCTTA
301 TGGGCGAACA CATGCTAGGA GCACAGGTCC CCGGCCAGTT GTGTATATGC
351 AGCATGCCCT GTTTGCAGAC AATGCCTACT GGCTTGAGAA TTATGCCAAT
401 GGAAGCCTTG GATTCTTCT AGCAGATGCA GGTTATGATG TATGGATGGG
451 AACAGTCGG GGAAACACTT GGTCAAGAAG ACACAAAACA CTCTCAGAGA
501 CAGATGAGAA ATTCTGGGCC TTTAGTTTTG ATGAAATGGC CAAATATGAT
551 CTCCCAGGAG TAATAGACTT CATTGTAAAT AAAACTGGTC AGGAGAAATT
601 GTATTTTATT GGACATTCAC TTGGCACTAC AATAGGGTTT GTAGCCTTTT
651 CCACCATGCC TGAAGTGGCA CAAAGAATCA AAATGAATTT TGCCTTGGGT
701 CCTACGATCT CATTCAAATA TCCCACGGGC ATTTTACCA GGTTTTTCT
751 ACTTCCAAAT TCCATAATCA AGGCTGTTTT TGGTACCAA GGTTTCTTTT
801 TAGAAGATAA GAAAACGAAG ATAGCTTCTA CAAAATCTG CAACAATAAG
851 ATACTCTGGT TGATATGTAG CGAATTTATG TCCTTATGGG CTGGATCCAA
901 CAAGAAAAAT ATGAATCAGA GTCGAATGGA TGTGTATATG TCACATGCTC
951 CCACTGGTTC ATCAGTACAC AACATTCTGC ATATAAAACA GCTTTACCAC
1001 TCTGATGAAT TCAGAGCTTA TGACTGGGGA AATGACGCTG ATAATATGAA
1051 ACATTACAAT CAGAGTCATC CCCCTATATA TGACCTGACT GCCATGAAAG
1101 TGCCTACTGC TATTTGGGCT GGTGGACATG ATGTCCTCGG AACACCCAG
1151 GATGTGGCCA GGATACTCCC TCAAATCAAG AGTCTTTCAT TAGTGCTAAG
1201 CCTATTGCCA GAATGGGAAC CCACCTTTGA TTTTGTCTGG GGCCTTGATG
1251 CCCCTCAACG GATGTTCACT GGAAATCATA ACCTTTAATG AAGGCATATT
1301 TCCTAAATGC CAATGCATTT TACCTTTTTC AATTAAAGG TTGGTTTCCA
1351 AAGCCCTTAC
(SEQ ID NO: 1)

```

FEATURES:

5'UTR: 1 - 100  
Start Codon: 101  
Stop Codon: 1286  
3'UTR: 1289

Homologous proteins:

Top 10 BLAST Hits:

CRA 18000004922653 /altid=gi 7434997 /def=pir  G01416 lysosomal...	431	e-120
CRA 18000004903706 /altid=gi 542751 /def=pir  S41408 lysosomal ...	430	e-119
CRA 18000004924799 /altid=gi 4557721 /def=ref NP_000226.1  lipa...	428	e-119
CRA 98000043616611 /altid=gi 12844223 /def=dbj BAB26283.1  (AK0...	415	e-115
CRA 98000043617058 /altid=gi 12845127 /def=dbj BAB26629.1  (AK0...	415	e-115
CRA 98000043616593 /altid=gi 12844194 /def=dbj BAB26272.1  (AK0...	414	e-115
CRA 98000043617174 /altid=gi 12845372 /def=dbj BAB26725.1  (AK0...	414	e-115

FIG.1A

CRA 98000043617140 /altid=gi 12845298 /def=dbj BAB26697.1  (AK0...	414	e-115
CRA 98000043617224 /altid=gi 12845477 /def=dbj BAB26766.1  (AK0...	414	e-114
CRA 98000043616955 /altid=gi 12844939 /def=dbj BAB26556.1  (AK0...	414	e-114

EST:

gi 8003062 /dataset=dbest /taxon=960...	62	4e-07
gi 8000757 /dataset=dbest /taxon=960...	54	9e-05

EXPRESSION INFORMATION FOR MODULATORY USE:

gi|8003062 Stomach normal  
gi|8000757 Stomach normal

Tissue expression:

Human leukocyte

1 MMWLLLTTC LICGTLNAGG FLDLENEVNP EWMNTSEII IYNGYPSEEY  
51 EVTTEDGYIL LVNRIPYGRT HARSTGPRPV VYMQHALFAD NAYWLENYAN  
101 GSLGFLADA GYDWMGNSR GNTWSRRHKT LSETDEKFWA FSFDEMAKYD  
151 LPGVIDFIVN KTGQEKLYFI GHSLGTTIGF VAFSTMPELA QRIKMNFALG  
201 PTISFKYPTG IFTRFFLLPN SIIKAVFGTK GFFLEDKGTK IASTKICNNK  
251 ILWLICSEFM SLWAGSNKKN MNQSRMDVYM SHAPTGSSVH NILHIKQLYH  
301 SDEFRAWDWG NDADNMKHYN QSHPPIDYLT AMKVPTAIWA GGHDVLGTPQ  
351 DVARILPQIK SLSLVLSLLP EWEPTDFFW GLDAPQRMFS GNHNL  
(SEQ ID NO: 2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

Number of matches: 5

- 1 35-38 NTSE
- 2 100-103 NGSL
- 3 160-163 NKTG
- 4 272-275 NQSR
- 5 320-323 NQSH

-----  
[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 4

- 1 125-127 SRR
- 2 204-206 SFK
- 3 243-245 STK
- 4 266-268 SNK

-----  
[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 8

- 1 53-56 TTED
- 2 130-133 TLSE
- 3 132-135 SETD
- 4 142-145 SFDE
- 5 162-165 TGQE
- 6 185-188 TMPE
- 7 274-277 SRMD
- 8 348-351 TPQD

-----  
[4] PDOC00007 PS00007 TYR\_PHOSPHO\_SITE

FIG. 2A

## Tyrosine kinase phosphorylation site

161-168 KTGQEKLY

[5] PDOC00008 PS00008 MYRISTYL  
 N-myristoylation site

Number of matches: 4

1	14-19	GTLNAG
2	117-122	GNSRGN
3	121-126	GNTWSR
4	175-180	GTTIGF

[6] PDOC00110 PS00120 LIPASE\_SER  
 Lipases, serine active site

167-176 LYFIGHSLGT

## Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	3	23	1.398	Certain
2	167	187	1.637	Certain
3	248	268	0.715	Putative

## BLAST Alignment to Top Hit:

>CRA|18000004903706 /altid=gi|542751 /def=pir||S41408 lysosomal acid  
 lipase (EC 3.1.1.-) /sterol esterase (EC 3.1.1.13)  
 precursor - human /org=human /taxon=9606 /dataset=nraa  
 /length=399  
 Length = 399

Score = 430 bits (1094), Expect = e-119

Identities = 211/394 (53%), Positives = 274/394 (68%), Gaps = 2/394 (0%)

Query: 2 MWLLTTTCLICGTLNAGGFLDLENEVNPEVWMNTSEIIYNGYPSEEYEVTTEDGYILL 61  
 M L CL+ TL++ G V+PE MN SEII Y G+PSEEY V TEDGYIL  
 Sbjct: 3 MRFLGLVCLVLWTLHSEGSGGKLTAVDPETNMNVSEIISYWGFSEEYLVETEDGYILC 62

Query: 62 VNRIPYGRTHARSTGPRPVVYMQHALFADNAYWLENYANGSLGFLLADAGYDVWMGNSRG 121  
 +NRIP+GR + GP+PV++QH L AD++ W+ N AN SLGF+LADAG+DVWMGNSRG  
 Sbjct: 63 LNRIPHGRKNHSDKGPKPVFLQHGLLADSSNWVTNLANSLSLGFILADAGFDVWMGNSRG 122

FIG. 2B

Query: 122 NTWSRRHKTLSSETDEKFWAFSFDMAKYDLPGVDFIVNKTGQEKLYFIGHSLGTTIGFV 181  
NTWSR+HKTLS + ++FWAFS+DEMAKYDLP I+FI+NKTGQE++Y++GHS GTTIGF+  
Sbjct: 123 NTWSRKHKTLVSQDEFWAFSYDEMAKYDLPASINFILNKTGQEQVYVGHVSQGTIGFI 182

Query: 182 AFSTMPELAQRIMNFALGPTISFKYPTGIFTRFFLLPNSIIKAVFGTKGFFLEDKKTKI 241  
AFS +PELA+RIKM FALGP S + T + LP+ +IK +FG K F + K  
Sbjct: 183 AFSQIPELAKRIKMFFALGPVASVAFCTSPMAKLGRLPDHLIKDLFGDKEFLPQSAFLKW 242

Query: 242 ASTKICNNKILWLICSEFMSLWAGSNKKNMNQRMDVYMSHAPTGSSVHNILHIKQLYHS 301  
T +C + IL +C L G N++N+N SR+DVY +H+P G+SV N+LH Q  
Sbjct: 243 LGTHVCTHVILKELCGNLCFLLCGFNERNLNMSRVDVYTTTHSPAGTSVQNMLHWSQAVKF 302

Query: 302 DEFRAWDWGNADNMKHYNQSHPPITYDLTAMKVPTAIWAGGHDVLGTPQDVARILPQIKS 361  
+F+A+DWG+ A N HYNQS+PP Y++ M VPTA+W+GGHD L DV +L QI +  
Sbjct: 303 QKFQAFDWGSSAKNYFHYNQSYPTYNVKDMLVPTAVWSGGHDWLADVDVNILLTQITN 362

Query: 362 LSLVLSLLPEWEPTFDVWGLDAPQRMFSGNHNL 395  
L S +PEWE DF+WGLDAP R+++ NL  
Sbjct: 363 LVFHES-IPEWE-HLDFIWGLDAPWRLYNKIINL 394 (SEQ ID NO: 4)

#### Hmmer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00561	alpha/beta hydrolase fold	46.7	2.5e-13	2

#### Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00561	1/2	112	195 ..	1	71 [.	38.8	6.7e-11
PF00561	2/2	294	352 ..	139	196 ..	8.0	0.19

```

1  TTATGGCCTA ACCTTTTAA CTTTGAGTTA TTTTCAAGAG AAAATTTGAA
51 AAAGCAGCCT TTGAGGAGAA AGAAGCAATC CAACAAACAA AAAGATAACC
101 AACTGTAAAT AGGAAATGTG TTTTGAATAG GACATTGGAA GAAAAATAAT
151 AATCATTTTT ACAGGTAGAT CCCAAAGTCA AGGATCTATG TTCAACCATG
201 TGTGTTCCAC CATCTTCACA ATTGAATGAG TAACCATCAT TAAGCAGTTA
251 GCTTAGGCCG TAATATGATT CTTGGACTGA GATTTCAAAA ATACCACAGG
301 CCTTCTGAAA GGTTACCCCT TTCTAGCTCC ACTATCATCT AATTTTATTA
351 AAAAAAAAAA AAAAGGAAAA ATTTGAGCTT CTAGAGAGTA GGGGCTACCA
401 TTTTGTATCC CACAGGGCCA AGGAACAAGT TTTAATGTAT TCATTTAAAT
451 TAATTTTCACT ATGAGTATTG AAATATATAA TAGAAATATT GTAACATTAT
501 ATATTTTCTA TATACTTTTA TTATATAGAA AATATATATT ACAGAAATATA
551 TTATTAAATA TTGTAGAACA ATATATAATA CAGAAAAATA TATAATACTC
601 AGTAATATAT TAAATACTTA TAAAAATAGC AAGCTTATAT AGGAAGAGTG
651 ATGGAGCATT GTGAGAAAGT TTCAGCTTTA TTTCTTTGAC ATTACTTTGT
701 TTCTGCACAA ACAAAGAAT TACAGGAATT GTCCAGATTA TTCAAATAAC
751 TCGAAGTTGA GGAGGGAATA TAAGTCAATG ATGTAGAAAC TCTTTTAAGA
801 TTTGAGCTAG CCTACAATCT GTAAAGATCT GTGAAATTGA ACTATATTTG
851 TGCTATTTCC ATATTAAGTC AAGGCAACAA ATCAATATTA ATAATAATAA
901 CATAGCACTT CTAGAACTTT CTAAAGAGTC CAATAAAGTT TTGTTAGAAA
951 GGATTGTTTT TGAAGTTAAA AACCATGAGA AATTCCAGGA AAATCCACAT
1001 ACCTATGCCA TCATACTATC AATCAGGGCA AAACATGCTT GAGTCTTTCA
1051 TCAAGACTAA ATGATTAAGG AGTGGTACAT AACTTTTCCC TGTTCTGACT
1101 AGCTGAACAC TTCCTTTTAC TCCACATTTG TTTAATTGGC ATGAAATTTT
1151 CCACTCCACT AAAACAGATC TTAGGATTTG GACAACACAA AATATCATTT
1201 GTTTTGAAAG GATTTGAGGA TAAATCCAAA CTAATAGAAC TGAAACTTCT
1251 ATATTATGCT GGGTAGCAAC TTAGTTTTCC CTACCCTTCT TCATGCTGGG
1301 AGATGAAAGA GATTCAGTTA CGGCTTAAGC TCCACAGGCA TACAAAGTGA
1351 AGCAGAAAAC TGAGGCACGT GTGCCTCCAT TATCTGGTAT CTCATGTGGG
1401 GCTTAGAGGT AAATTGTCGT TATTTGGCCT CCATTTCTGC CTTTAACCAC
1451 TGGTGTAAC AAAGTTACT GTGCCAAAGT TGACAGCAAC CCAAATCCCT
1501 TTGGCATGTG AATTAGTTTC CTCTGCCATA CTGCTAGTTC CAAATTCCTT
1551 CTGGTTTCAG GATTTAGGAG TCAGGGTTGC CTCATCTTCT CAAATGAGTT
1601 ACAGTCACGC ACATCCCTAC AACTGTCATG GTTGGCACTA GTTCCTTGAT
1651 ATATGTTACT CCGTTTGATC CTCATGAAGG ATCAAATGGG GAAGGGAGAT
1701 ACTATTGTCT CTGATTGTCC ATTAAGATCT TGAGTATGTT CTACTTCCCT
1751 GTTTGACACA CTGGTTTGAA AATGTTGCTA AGTCTTCCCA ACAATGACAG
1801 AACTCAGTG GAAACATGAA GGATTCCGTC AAAGTGGTTA TTTTGCATCA
1851 TGAGACCAC TATTTCCCAA CCTGCAAGTG CATCATGGCC TTTGGTGTGT
1901 CAGGGACACG CCTTGGGTGT GTGTCTCAGT CTAAAGCTTC CTCCTTTTCA
1951 CAAGCTTCCT GTTTCTCATC TCTCTAGCTT CTAAGTGTCA CTGTAATCAT
2001 CTCTTACTCT TCAGCCTGAT GTCAAAAGCA AAAGTTCAGA AGTTCCTCAT
2051 CAATAAGGAG TCCTTGAGAG CAGGTGAAGC TCATCTAACT AGGTAAGATG
2101 AAGATCTATC ATAACCAGGA GGCAGGTTGG AAGGTGCCAG TTGCACTGGC
2151 AGTCAGGTGC AAGAGCTCTG CAGTGAGGCT GCCTGAGTGT CCATCCTAGA
2201 TCTCTCACCT CTTGGCTCTG TGACCTTGAG CAGGTCTTAA ATCTCTCTAA
2251 GCCTTTGTTT TTTAATTGA TAAATGAGG ATAATAATAG TACCAAAT

```

FIG. 3A

2301 AGGGAGATTT TCAGAGCTTA AATAACATAC GTGAAC TATT TAGAGTAATG  
2351 CCTGCCATAA GGGGACTCAG TAGCTTATTA TTAGTTTCAT ACAATTTGAA  
2401 AAGTTTCATA ATATTTGCAG ATATAAGATG ATCTTCAACC AGATAGCTAA  
2451 TGTATGCAAA GCTATTTAGC TTCAGAAGTA AACTCTGCAT TTCTAGAAGT  
2501 TAAATATTAC TTTGTTATAG TGAATTATCT GTAATATTTA TCTCTTGCTC  
2551 ACTTTTATAA GAAAAATAGT GAAAGCATTT ATTAAGAACT TACACTGCAC  
2601 TAAATGTTAT ATATGACTTA ATCCTCACTA TAACCCTATG AGATAGGTTA  
2651 CATTATTGTC CTAATTTTAC TAACAAGGAA ACCAAGAGAC AAAGCTACTA  
2701 AAACACTTGC CTGAGGTTAG ACATCTTCTT CTGTGGTGAG GCTGGATTTC  
2751 AAATTTAGAC CATTTGACTG TAGCACTTAT ATGATGAGCA TGCTGTTTAG  
2801 TGTTATAGTG TTGGTCTACC TTTGAATAGA CATACTTTTA AACCATGGCA  
2851 AGGAAGTGAG ACTGCACATT GAAATATGTA AAATTTGCCT TTGGGTGCCA  
2901 CGTGAGAAAT AGTCACATCA CTAGAACTA ATCATAAGCT TTTGTGTTTG  
2951 GTTAAAGTTT TATTGATCCA TTTTCTTGT TTACTTTGTG GGATACTGGG  
3001 CTTAACTAGG GGATACCTCC ACTTTTTACT TGGCCATGGT ATGAAAACCT  
3051 GTCCTCTGAA TCTTTAGATA TTTTGGCAA TGTAGGCAA ACAAAGACTT  
3101 AAAGCAATTC AACCTTGATT AAAATAAGAC CAAAAATGCC TCCATACTTG  
3151 ATTAAATTTA TTTCAATTTA GGAAGTGGAT TATAATCAAG ACAACTTCTA  
3201 CATGAAAAAA TAGATTAATA GTGCTCCAAG TAGTTCACT GTATTTATTC  
3251 CTTTTTATAC ATTATCTGCC TTCGGTGTTA TTCAAGTTTT CATTAAATCAT  
3301 TAATAATTTT ACTAATCATT TTATTTTCAAT AATCAACATT GATAGTTAAA  
3351 ATTAATCTGT GAATATTAAT TGTTTTATGC CAGGCATTTT TATGATGTGG  
3401 CTGCTTTTAA CAACAACCTG TTTGATCTGT GGAACCTTAA ATGCTGGTGG  
3451 ATTCCTTGAT TTGGAAAATG AAGTGAATCC TGAGGTGTGG ATGAATACTG  
3501 TAAGTCATGG AAAACTGTGA AGAACATCAA ATAAAGCAGG ACTAATGGAG  
3551 TATGAGGTTA CGAAAGGTCC TGTGTGAACA GAAAATCTCT GATAAAACAG  
3601 ATAAAAATGA GATGGTTTTT AACCTCTGCA AGAGTCAAGC TAGTTAGATC  
3651 TTTGTCTGAA AAACAAATAC TGTCGGGTAA TGAAAACCAA ATTGTGCTAT  
3701 TGTGCTATCT ATCTATCTAT CTATCTATCT ATCTATCTAT CTATCTATCT  
3751 ATCTATCTAT TTATCTATCT ATCTATAGAT AGAACCTCCT CTTTTGAATT  
3801 TATGTTTTAA GAATATCAAG CTATTTGTTG ATATACATGA TTGCCTTCTA  
3851 TTGATCTATA GTTCTATTAC TTTTAAAGCA AGAGGGGTCT CAAAAGACAA  
3901 TTGACTTGAT AATATAGCTT TGTCAGAAAG AATGGGTCAA TGCTAAATTT  
3951 TCCCCCAACC CCCCCAAATA TTAGCCAATA GTAGATATTT TTTAAAATTC  
4001 TACTTATTTT GTATTAAGAC TTTATTTATT AATTTTACAG TTACCTGGTG  
4051 CTACAAATTT CAGATAATTC ACCCTAATAA GCACACAACA GATGGTTTGT  
4101 TTTGATTCCT TTTTATATCC TTTGGAGAAG TTCCACTAAC GACTGTATTT  
4151 TTAAGTGGCA GAGTGAAATC ATCATCTACA ATGGCTACCC CAGTGAAGAG  
4201 TATGAAGTCA CCACTGAAGA TGGGTATATA CTCCTTGTC AACAATTTCC  
4251 TTATGGGCGA ACACATGCTA GGAGCACAGG TACAAGATAT GTCTCTCCTG  
4301 AAAAGGGGAC TGCATTGACC TCCTGCTTCT CAGGAGGAAT TTAATGCTAG  
4351 ATATGCATCA ACAGAGTTTA TCAAAATTGG TTTGAATTAT TGGATTAGTC  
4401 TTTAAATAGT TATCAGGGAG GCTCACTCTT TGCCTGATAA TTCTCTGAAG  
4451 ACAGACAGGA ACCTAAAAAT ACAAACAGCA AGACTGATCT TGCTAACTGC  
4501 AACCAGAGGT ACTTGTTAGG GTGTAAACAG AAAGGCAGAG CCTGCATTTT  
4551 GTCACCTCAT TACTGATTTA TCATGTGGAA AATTGCTTTG TCCCAGGAAA

FIG. 3B

4601 ATGGATCCTC TCATTGTCAG AAGGAGATTT TCTAGGTTGT ATGAAATTGA  
4651 CTCTGGGGCA CCCAAGAAGA ACCTCTCCTG CTCCCCTAA AATTAAGGGG  
4701 CCTCCCTCTG CAGGATAAAA AACAACTAG TTAAATGACA ACGCATTTCT  
4751 GAAAAGTTTT CCAGGACTGA AAACCTTAAC ATCCACATAC ACTTTGATCT  
4801 AAGGGACAGA CGGTTTCATAG AATGAAAGAG TATGGTGTCA ATAAGGCTTG  
4851 AATTCTAGAA TGAGGAGCCA GCCATGCCAT AGCAGGGGAA TGATACTCCT  
4901 TAAAAGGGAA AATTTAACTA CAAATCCTCT GAAGTAGAAA TGATAAGAAT  
4951 AACCAAAATA TCTGCAATGG TTCAATAGCA AATAATTTAT TGGCAGCTGC  
5001 TTACCGTGTT CATTTTGCAT CTTTTTCCC ACCACACATA TTAAGGAGCA  
5051 GCTGAAGTCA TGTTTGACAT TCTCTCCCTC TTTTATCTCC AGTTTCAGAA  
5101 TGAAAAATGA GAGTGAGATA TGAGTAGTTT TACTAGTTAA AATATGAAAC  
5151 ACCCAGTTAA ATTTGAAGGT CAGATAAACA ACAAATAATT TTGTATAAGT  
5201 CTCATTTTAA GATAATACTA AAAAGTCATT ATTTATTCAC TATTATCACT  
5251 ATTTATAAAA TTTTGTAGAG CATCCTGGAT CTTTTTGCTT ACTTTTGTTC  
5301 TTATTTTTTG CTAAATCTGG CAATCCCAGG CACATGTGTG AAGGAGCTGT  
5351 GAAATATAAA AGGAGAAAAC TTTTATGGGA AAGATTTGGC TTAAGGAGAG  
5401 ATAATTTTGG AAAGATTTAG AATTAAGAT CATTCAATAG ATGTAATGTT  
5451 CTAAATACTT TATATCAGTT AAACCTTCTCA TCAACAATAT GAGATGGGTA  
5501 CCACTAATAG TCACCATTTC ACAAATGATG AAATTAAGGC ACAACCGGTT  
5551 ATGTTAAGAG GCCTAAAGTC CACAAATAGC AAGCTGACAG ACCAGAATTT  
5601 AAGCCCAGGC ATGCTGGCTC CAGAGCCTGT GCTCTTAGTC ATTAAATTAT  
5651 AGTGCCTTAC TTGACCTTCC ACCCTGGTTA CTTTGGATCT CCCTGAATGC  
5701 TCTCTCTCCC TCAGAAATAC TGGAAAGTTGG CAGAGGGACA CTGAGCTGAG  
5751 CATATTATTG TAGTTTTTAA ATGCTCTCCA CTGGACAGAA GATGGGGGAT  
5801 TTGAATAGAA ATTTGGTGAG GAACTAATCA GTGTCCATTT ACACTCACCT  
5851 CCTCTTCCTC CCTGGAAGAG CTATAGGACT TGAGTAAGCA TGATAAATTT  
5901 CGTGTCTTTG TAAACCACAC CCAGGAAATT TGTATATACA AATACATAGA  
5951 GCACAGTAGT TATCAGGACA GACTTTGACA TAAAAAGAAC TGGGTTTGAG  
6001 TCCCTGCTCT GGCCTTCTTA TCTGGGTGGC CCTCTGGGAA AGTTACTTAA  
6051 CTACATAAAG TTTTGTTCCT ATATCTACAA AATGAGGTTT CTCAAAATAG  
6101 CAGCTAGTTT ATAGAGTTGT TGCAAGAATT TAGTAAGCTA ATACATATAA  
6151 ATACGTCAAC ATAGCACCAG GTACAAAAAT ATGTGCTCAA GAAACTGAAG  
6201 TTACCTGATT ATAATGCTCT ATACTATTGA CAAGGGAAAA GTGAAAACAG  
6251 TTTTGTGTTT ACCATGTGTG TATGTGTGTG TGTCTGTGAT GTTTCCGACA  
6301 TGCTCTATTT AACATAAATT ACTCTCACTC TTTCTCTCTC TCTCTTTCTC  
6351 TTTCTCCCTC TCTCATCTTA CCCTTTCCCC CACCAGGTCC CCGGCCAGTT  
6401 GTGTATATGC AGCATGCCCT GTTTGCAGAC AATGCCTACT GGCTTGAGAA  
6451 TTATGCCAAT GGAAGCCTTG GATTCCTTCT AGCAGATGCA GGTTATGATG  
6501 TATGGATGGG AAACAGTCGG GGAAACACTT GGTCAAGAAG ACACAAAACA  
6551 CTCTCAGAGA CAGATGAGAA ATTCTGGGCC TTAGGTAAA TATTAGCTAA  
6601 GAAAACCTCA GGGGGAAATT GGAGGCAATT TAAAAAAAT AACGTGGACG  
6651 CTATTAATGA TTATCTTTGA CGTTTGAAGT CATATAGCTC CTTGTAGTTT  
6701 CTGTTAAGAT CTCAAAGGAG GGTAACAGCA AGAAGCTCTG ATTTTTCAC  
6751 GATTCTCCCA CAAGCAAAGT ATGGCATTTC AACAAGATCA TTTTACATC  
6801 CAATTCTGTG AATTCTATGC ATTAAGAGTA TGTCAAAGA GACAGCTCAG  
6851 GAAATTATCA TGACCAATGT GCACATTCAT TCAGCCAATG TTTACTGAGT

FIG. 3C



```

6901 GGCTACTGTA TGCCTGTTC TAGGCCCCGA ACATTCAAAC AGGGAACAGA
6951 CAAACTCTGA CCTCACAAG CTTATGTTCA TTTTAGTGAT AATTTTACAA
7001 GTCATTGCTC CTGGATTGCC AATCAACTGT GTAAAGATGA TTTGGACCAG
7051 GACCTTATTG ATTTAGAGAA ACTGTGATTG ATTTAGAGAA ACTGAGATCG
7101 CACATAGTAC CATTTTCAGG AAAACTCCAA TATTAGATTT TAAAAACCTT
7151 GTTAATGGGC AATGAAGAAG AATCTTTTTT GATATCTTGT TTCTTTTAAT
7201 GGAAGAGTTT TCTGCTGTCA CCAGAGGACA GGCTGATGCC TGCGATAGAC
7251 TTTTCTTTCT TCAGGCCTAA GCTCCCTGTT GGTTTGTAAG CCTGATGCTA
7301 GAACAGACTG TGTATTCCTA TTACATTAAT AAAACATTCA GTACCCACTG
7351 AAAGTTTGAG AATAGTGGAG GAATAGAATA GAATGTTATA GTCTGAGTTC
7401 TTGGGCAGGG GCAAGCATCA GGAAATATTG AATCATTAGT CTTTAGGAGG
7451 TGTACAACA ATTCTCCTAT TCTTGTAAGT CCAATCTAT AGATTTCTCT
7501 ACATGTTCTT TTAATAAACA GGCTTCTAGC TTATGGAATA CCTGATTTGA
7551 CTAAATGTTA TATAGGCCCT TTTGTTCTCT CTGTCTGAAG AACAAAATAC
7601 TAGTACTATG GAATATTGGT ATATATTAAG TATATATCTA TATATCCATG
7651 TGGACAGGAA TACTACTACT AACACATCT TACTGAGCAC CCACTGGCAG
7701 CCAGAGTCGT TTCTTTCATA CTATTAAACC CCGTTAGCAG CCCCCTAAAC
7751 CAGGTACTAC CCTGTTTATT TCCCAAATGA GAAAACATAG GCTCAGAGCA
7801 TTTCAGTAAT TTCTCAAGAG TTGCAAAGGC CATAAATAGT AGAATCATGA
7851 TTTACAAAAC CCCTGTTTCC AAAGATGGGT ATTAAATGGT CCTAACAATT
7901 GTGAAGCCTC ATGTGGGAGT CAGAAGTAGA GGCACACAAG CCAGATGGGG
7951 AAAGGGAGGG CAAAGAAAAG CAAGAGAAGG GAAGGAAGAG GAGGGATCAT
8001 AAGGTTGAAC TTCAAATATC ATACACAAGT TTCGAAAGTG TTCCTCTTAT
8051 AAGGAAGTAA AATGTACATA TGCAGAAAAA CAAAAGCTA CAATAGCCTA
8101 CATATAATTG GATAAATAAT GAAATACACA TTGAATCTAA GTAAACAGCA
8151 TAGAATCTGG GTGTAAAAAA GAAGTGAGCA AGTGCTCTGA GTTTTAAACT
8201 TAAACTTGCA AGTATTTATA AAAGCCCCTG TTTTATTTTG CAGTTTGTAT
8251 GAAATGGCCA AATATGATCT CCCAGGAGTA ATAGACTTCA TTGTAAATAA
8301 AACTGGTCAG GAGAAATTGT ATTTCAATTG ACATTCACCT GGCCTACAA
8351 TAGGTATGTT TATGAGGGTC ACTGTTAGGT GTGTTTTTGA GGGTCAGTTT
8401 TCTCAGAGTC TTACAGGAGT TCACCTTTAT GTTGAATAA AACAACTGTT
8451 ACTTATAGTG CCCTCAATTC CCTGTCCTCT GCTGGAATA ACCCTAGTAC
8501 TCTAAGTAGC TGTGAGCCTG CAGTGACAG ACTATATGTA GGGCAAACCT
8551 TTCCTGGGTC TCTGGTCACA GCAGCATATT GACTACGGTG ATGCAATTC
8601 CCAGGAATAA CATGTGTTCC AAATTCAAAG AAATAATTCC ACAGAGTAAG
8651 TTTCTAGATT CCCTCTGAGC TGAAAAAGTA AAATTCAATG CCATGGAATA
8701 TGGCTGAAAC ATAATAAATG TGCATCAATC ATCTCTTTCT CACAACCCAA
8751 ATGGGATTTT TAAAAAATAA AAGGGAAGGG CTTATACCTA TATTTAAACA
8801 AATTGAAAAG GCATGGTTAT ATTTGTTTGT GAGTTGGAAC ACACAAGCTT
8851 ACTATAATAA ATCAATTGAG CTTATCTATT CAGTGTGTGA TTTAGTATTT
8901 ATGAAATAGC AAGTAAATGT AAGCACTATG TAGAAATTTT TAAAGTTTTT
8951 TAAGCTGACA ACTTACTTCT TAATTTACTT ACTTTACTTA ATTTACTTTA
9001 CAATTTACTT TCCAGGTATT TTGGAAAGAA ATCAATAATC TAGTTCCAAG
9051 TAAAAGTTGA AAGGAACCCA CACTAATAAA AGCTTTGAAT TTGTCATTGA
9101 ACTTCCACTA AAGTTTCCAA TTTTAAGAGA ATAAATCATG TGAAAGTGCA
9151 ATATTTCACT TTAGGGAAAT ATTTTCATTA TCACCACTAT CATCAGTAAC

```

FIG. 3D

9201 AAACATATAT TCATTAGTAT TTTAGATTGA CAGGCACTTT CCAAGCTCAG  
9251 AACAGGCAGT TAGCATCAGT CAGCATATAC TAAAAAAGTA TCAAAGAACT  
9301 CATAGGAGAT CAAAAATGCC ACCAATAGGC AAATAATTAC AGTATCTAAC  
9351 ACTTATTGAG CATTCGTTAT GTGTAGGGTC TTGTGTTCAG GACCTTCCCC  
9401 ACAGTATCTC CCTCTGATCT TCAAAACAAC CCGAATGTTA TTATCCCCAT  
9451 CTCATAGAAG AAGAAACACA AGTTCAGAAC ACAGATTCAA ACCAGATGTA  
9501 TCTGATTTCA CCAATAGGGT GTGTAAGGAT TCCGGAGAAA TGGTGTAGAG  
9551 AAGAAGAAAT GACTTTAGTT GGTTTTGGAA AGTGGGTAGG ACTTAGATAT  
9601 GCTCTTATAC TTGATCTGCA AAAAAAAAAA AAAAAACCAT GGAGAATTTG  
9651 ATTATCTGTG CTCTGTGTTT CATTTAGGAC ATAAATATTT TTAGTGACTG  
9701 TTGTTTGCAT TTTGGACAGA GCAATTTCTG TTATGTAAGG AGCACCCACT  
9751 CTTTGTAGGA CATTTAGTAG GTCCCAGCCC ATTAAACAGG GCTCTGCAGT  
9801 CAGCGTGACC CTCAAAAATC TCACCTCCAC ACATTTCCAA ACACCCTCTG  
9851 GGGAAAGTACT ATTCCTGATT CAGAGTCTTT TTATCAATTG TTCAGTCAAT  
9901 TATTTCAAGT CTCTTTTTTC TGGCCAAGAC AGTTTTAATG TTCCAACAAG  
9951 TGTTTCAGTA CACACATACA CACACACACA CACACACACA CACACACACA  
10001 CACATGCTAG TGGAGGCCCA GGAAGGGACC TCTGGAAACC AAATTATATG  
10051 GATATTCTCC CTAGCCTACC CAGTGTGTG CTAATCTCCA TCCTCACAGA  
10101 TATACAAAGG GGTGCAATGC TACTGCTGAA AGAGCAAAGC AAATGGAGAT  
10151 GCCTGGTCCT TACTGGGCCA TCGTGGATGC TAGGGAAAGC CCCTTTCTTT  
10201 TTGGAAACAG GGAAGAGTCT AGAGGGTTGA AAAACACCCA GTAAGACACT  
10251 GGGAGCAGTG AAATTTCAAT CCATAGTGAG AAAGAAAACC TGTTAGAATA  
10301 ACTGGGTGAT GCTGCAGAAA GAAATCAATT CACCTCCTGT GACTGATTAT  
10351 TTGCTTCTGG AAGCTCTGTG ATTCATTCTG GCATCTCAGA GTTAGGGATG  
10401 AAATGAGAAT GTTGCCAGCA TTTACCCCAT GCTTGGGAAG TTTACACAGC  
10451 AGTAGCTACT CCAGCAGCTT AACCATCACC TTTCCCCTGC CAACTACTCC  
10501 ATTTCCCCCA ATCAAGTCAA ACTGTCCATA AATAGAATAA AATAAAATTG  
10551 GAGACTTGAG AGCAGAGAAG ACTGAAGGCA GATTATCTTT ATAGAATAAC  
10601 TCAGAAGACT TCCAATTCAT CCCCAGTATG ATCACGATAG AAGGAAAAAA  
10651 TGACTAAGCA GAGCCCCAAT TTTGTTAGAA ACATTGCGTA AGTATTTATT  
10701 TTTACAAGAT TGTCTTATCT CCTGTTCTCT CAGGGTTTGT AGCCTTTTCC  
10751 ACCATGCCTG AACTGGCACA AAGAATCAAA ATGAATTTTG CCTTGGGTCC  
10801 TACGATCTCA TTCAAATATC CCACGGGCAT TTTTACCAGG TTTTTTCTAC  
10851 TTCCAAATTC CATAATCAAG GTAGGCTCCT TTCAACAAAA TGTAACCTGAG  
10901 GATCTCATTT TGGATCATAA ATCCTTATTA TTTTCAAATC TACTGTAAAG  
10951 TAAAAGTAGG AAATTTAGAT AAAATCTATA GAACTTAGAC TCTGTGGGTA  
11001 TGTGCTTGTG TATGTGTGTC CCTGCGTGTG CGCATGTCTG TGCCATAGTA  
11051 TCTGCAGGTT CTGTAATACA ATTTACTATA CAAGGTCATC AGCAGGCTGA  
11101 GTATATGTCA GAATTTCTAG CTGAACTGAG TGCTATATGA CAACAAGGAT  
11151 TTTTCTTGTT TTCCCAAGTG TTTTTTGTTT CATTTAGTCA GGTAGGTCAA  
11201 TGAATTCACA TTGCCCAAAT GAAAGACACT TCAAGTTACC CATAATCACT  
11251 GATGTGTCCA ATTTTGACAT TAGAAAAACC TGATTAATAT ATTCCTTCCA  
11301 ATATGGAAAC TTGCCCTAAT AACTAAAGCT AAGATTCCAA AGCCTAAATG  
11351 TATTACAGCT CAAGTATTAA TTCAAATATT TATTGGTTAT TTTTCAGGAG  
11401 TTGAAAAAGT CATTTGGTTG CCAATTGTGG ATTTGGGATT TTATCTATTA  
11451 AAGGGTTTTT TTTTTTTTTT TCTTTGCTTT TGTTTCTCTA CAAAGGTCAT

FIG. 3E

```

11501 TGCCACAATG AACACAGCAT TTAATCAAAT TCCAGATTGG CCTTTGAACT
11551 TGGGATGATG GATAAAATGG ATTTGGGCCA AAATTGAAGT CAAGGAGACC
11601 AGTTAGAATA TCAAAATAAT TCATATATAA GAAAATGAGA CGTTGGTTTG
11651 GGGTAGAGTG GTAGGAATGA AAAAAATTAT TTGTGAGCTA ACACAAGGAA
11701 TAATTTCCAT AGGGCCTAAT AATAGTTAGG TCTGATAATA CTATGGTCTG
11751 ATAATAGTTT TATTGTATTG TTTACTGAGA GCACAAATGA TGTAACCTCC
11801 TTATTCAAGA GCTTTTCTAG TTTATTTAAA AATGTGTTGA CATCAGTTAG
11851 GTTTTAATGT TTTCTATATT TGGACAGTGT GAGCAAACCTA ATTTGTAAAA
11901 TTAAATTCAG AGAGAGATAC ATCTATCTGT AAATACATAT ATGCGTTGTT
11951 TGTGTTGCTC TTCCTACATA GGTGAGCTAT AAGGCAAATA ATGTTCCCTGG
12001 GTTATCTCAG TTTCACATTT CCCACTGTCA ATATTCCTGC TACTTTTAAG
12051 TCCCATATCC TGCTCTTTTC TTCCGTCAGT TTCCCCCAGA AGCTCCAAGA
12101 CCCCACCAGG AATCCCCATC CAAGTTTACT TTCCCAACTC CTGGAAGTTT
12151 CAATTGTGCT GCCTTTGTGA CATTATCATA TCTTTTCTGT TCAATGGTTG
12201 CTTCTCTTTG GCTCACTGTT CTCTACTTTT CAGCCTGAGA GCTGGCTAAT
12251 CTGGGACAGT ACTCGAATGC AGTGTACACA TGGGTAACAT GGAAAACCCC
12301 GATTTTCCCT TATATTCAAG GTATTATTTG ACCTTAAGAA AAACTGTTTT
12351 ACATTTTATA CCAATTAATG AGAAAAAAAT ATTGGCAAGC ACTGACTGGG
12401 CAGAATACAG GGAAGCTTCA CTATGGAGAA GTGAATTTGG GATTGAGGGC
12451 CTTTATTGCA ATCTCCTTGT AAATAATATT TGATACTCTT CCTCATCTGG
12501 AGACACATTC CTAAGTAACT TTTCTGAAT AATTTGGTCT CTTGACTGA
12551 ATCAGTAAGT ACAAATAGAT CCCCAGCAT GGCTCTTTCC TAGAATGAAA
12601 GAAATGTCAA GAAGTCTGAA GATGATTCTT GAATTTTGGT TTTTGTCTAT
12651 TGCTATTTGG GCTTGTGTGC CTTGTTGTG CTATTGAGTT GAGCTCCTTA
12701 TATATTCTGG TTAATAATCC CTTGTAATAT GGATAGTCTG CAAATATTTT
12751 ATCTCATTCA AAGATAATTA TTATTTACTT TCATAGGCTG TTTTGGTAC
12801 CAAAGGTTTC TTTTGAAG ATAAGAAAAC GAAGATAGCT TCTACCAAAA
12851 TCTGCAACAA TAAGATACTC TGGTTGATAT GTAGCGAATT TATGTCCTTA
12901 TGGGCTGGAT CCAACAAGAA AAATATGAAT CAGGTATGTA TGATAATTAT
12951 AGGGCCATTT GATACCTTAA GAAATCCAG CTTTCCTTTG ACTCATTTTG
13001 ATATATCTAT TTAAGTATA AATTCATATG GTATTCCAAA CCCTTAAAGA
13051 CAGATTTTTT TTTGCTTTTA AAAATGTTTA TGGGTATATA ATAGTTGTAC
13101 ATATTTATGA GACACATATA TTTTGATATA AGCATACAAT GTGTAATGAC
13151 CAAATCAGGG TAATTGGGAT ATCCATCACC TCAAGCATT ATCATTCTT
13201 TTTGTTAGAG ACATTCTAAT TTGACTCTTC TAGTTATTTT GAAATATACA
13251 ATGAATTATT GTTAACTATA GTCATCCTAT TGTGCATGCC AGACTTTAGT
13301 CTTTCTAACG GTATTTTGGT ACCCATTAAC CAATGCCTCT TTATCCTTCC
13351 CCCACCCCTA CTACCTTTCC CAGCCTCTGG TAACCATCAT TCTTCTCACT
13401 ATCTCTATAA GGTGAGTTT TTTTAAACT CCCCTATATG AGTGAGAACA
13451 TGCAGTATTT GTCTTTTGT GCCTGGCTTA TTCACTTAA TGTAATGTTC
13501 TCTAATTTCA TCCACATTAT TGCAAATGAC ATGATTTTCT TCTTCTTATG
13551 GCTGTCTATA TGTACCACAT TTTATTTATC CACTCATCTG TTGATGGACA
13601 CTTAGGCTGA TTTCATATCT TGGTCATTGT GAATAGTGCT GACTAAACA
13651 TGGGGGTGCA GATGTCTCTT CCATGGATTG ATTTCTTTT TTTTCTGTA
13701 ATATAGACCT AGCACTGGAA TTGCTGGATC ATATGGTAAT TCTACTTTTA
13751 GTTTTTTGAG GATCCCTCAT ACTCTCCCC ATAGTTCCTG TACTAATTTA

```

FIG. 3F

13801 CATTCTACC AACAGTCTGT GCAAGAGTTC TCTTTTCTCC ACATTCTTGT  
13851 CAGCATCCAT TATTGCCTAT CTTTTTGATA AAAGCTATTT TAACTGGAGT  
13901 GAGATAGTAC TTCATTGTAG TTTTAGTTCG CATTCTCTA ATGATTAGTA  
13951 ATGTTGAACA TTGTTTTTAA TGTACCTCTT GGCTATTTGT ATGTCTTCTT  
14001 TTGAGAAATG TCTACTCAGA TCTTTTGTCC ATTTTTAAAT CAGATTTTTT  
14051 TTTTGCAATT GAGTTATATG ACCTCTTTAT ATATTCTGGT TACTAATCCC  
14101 TTGTCAGATG GGTAGTTTAC AAATATTTTC TCTCATTCAA CAGGTTCTTT  
14151 AGTTCACTTT GTTGATGGTC TCCTTTGCTT TGCAGAAGCT TTTTAGCTTG  
14201 ACGTAATCTA ATTTGTTTAT GTTTGCTTTG GTTGCCTGTG CATTGAGGG  
14251 CTTACCTCAA ATTGGCCCAG ACCAATGTCC CGGAGTGCTT CTGTAATGTT  
14301 TGTTTTTTAG TAGTTTCATA GTTTTAGGTC TTAAATGTGT CTTAATCCA  
14351 TTTTGATTTT GTTTTTGTAT CTGGCAAGAG ATAGAGATCT AATTTTCATC  
14401 TTCTGCATAT GGATATCTAG TTTTCCCAGC ATCATTCTT GTGGAAATTG  
14451 TCCTTTGCCC AATGTATGTT CTGATGCCT TTGTTGAAAA TTAGTTGACT  
14501 ATAAATGTGT GGATTTATTT GTGGGTCTT TATTCTGTT CATTGGTCTA  
14551 TGTGTCTGTT TTTATGCCAG TATCATGCAG TTTTGATTAT TACAGGTTTG  
14601 TAGTATAATT TGAAGTCAGG TCATGTGATG CCTCCAGCTT TGTTCTTTTT  
14651 TCTCAGAATC TTATATTTAG AAAAACGTAA AGACTCCAAC AAAAAACCTG  
14701 CTAGAACTGA TAAACAAATT CATTAAATTT GCAGGATACA ACATCAACAT  
14751 ACAAATTC A GCAGCATTTT AATATGCCAA GAGCAAATAA TCTTAAAAAA  
14801 AAGAAAGAAA AAAAAACAAG AAATAATCCC ATTTATAATA GCTACAAATA  
14851 AAATAAAACA CCTAGGAATA AACCATACCA AAGAAGTGAA AGATTTCTAC  
14901 AATGAAAAC ATAAAAACAT GATGAAAGAA ATTGAAAATG ACATTAATAA  
14951 ATGGAAAGGT ATTCCATGTT CATGGATTGC AAGAATCAAT ATTGTTAAAA  
15001 TGTCCATATG ATCCAAAACA ATCTACAGAT TCAATGCAAT CCCTATCAAA  
15051 ATACCAATGA CATTCTTCAT TGAAATAAAA AAAAAGCCTA AAATTTAAGT  
15101 GGAACCATGA AGGTAGATGT CTGCTATACA TAGAAGATTA AGTACTCAAC  
15151 AAACCTTGAA TATGAAGACT GGGGAAGTGA ATAGGCAGCT TCACTCTTCT  
15201 ATTCCCTGGT GAAATTTAGG AGAATGGATG TTTTATAATG GGTAGCAGTT  
15251 TCTTACATGT TCTCAATCAG CCATAACTTA CTACAGTCAA TTTGAATTTA  
15301 TTGCATTTGA ATATATTGGA TTAATAATAA AATCCTAAAA AAGGAGAGAA  
15351 GCACATATAA ACCTGCGTCT TATTTTATGT GTTCCTTTCT TTGTGGGTGA  
15401 CTTTTGTTTT GAAATAAAAC CTGCAAAATA ACAGGACAGG GTGGAAGGGA  
15451 GATGGGATCC CCTCTTTATG AAGAAGCAGC AGTCCTGTTT TATCACCTCT  
15501 TCATTTTCTG TTATTGAGAA TTCAAGAAGA AGGAGGAGGA AGAGTTCACA  
15551 TCCACAGACT GGTGTGGTTG AATAGTTGTC TCTACTGTAT TCCAAATAGC  
15601 AGCCAATGAG GCTGTTACAG TGAAGCCAGT CCCAAGATAA TTGTTCTGTA  
15651 CCCCTATTCT CTAAGAAGCT AAATTGTGTT AGACTGAAAC CCATAAGGAA  
15701 CCATTGTTCA AAGTTGGCTT GTTCAAAAGT AAAGATTTTT AATAGTTTCT  
15751 CTTAATTAGA TTATTTTCTA AGACATAGAA TTATGATTAC TATTTTATCT  
15801 CTATAATTTT CATCTCTATA ACGTTTACAA ATACTGAAAT AACCTTTGGA  
15851 AAAAATTGGC TTTTAGCTTT ACTTTTGCAA TATTTTATTT TATCCCCATA  
15901 AAAGCCTAGG AAATTGGTAC TATGACTTTT AGTATGTTCA TTTAATAGAT  
15951 GAAAACACAG AAACCTCAAAG ATGTTAAATA TGGTGGCCAA GTTCACAAAG  
16001 CTGATCATT AACAACAACAG GGCCTGAACT CCTGGTTTTT TGATTTAATC  
16051 TGTGACAGTG CACCTGGGTG CGCATGCATG CATCACCCCC ACATTGAC

FIG. 3G

```

16101 ATAGAACCTT TCCTAGTTGG CTTTGCTCCA TGATGACCAT TACTGTTCCCT
16151 TCTACTTCAA AATAAGCAAA TTATCCTACA GATTCAGAGC TGGTACAGGT
16201 GTGCTGTCAA GCAGCCCAT CCATTAGTCA GCTTGTGGTT CACTCACATT
16251 AAAGTATTGA CCTAAATGGT ATATTTATCT AGATAATTCT ACCTTGTTAT
16301 TTTCAAAGCC CCAGTCTTGT TTGCTAATTC TGTGCATCAT TTTTCTCTGA
16351 TTCTGAAAGG CAAAATTTTG TTGGGCAATT GCTGTAATAT GAGTTTTATC
16401 TCCTTTAGAG TCGAATGGAT GTGTATATGT CACATGCTCC CACTGGTTCA
16451 TCAGTACACA ACATTCTGCA TATAAACAG GTAGAGTCTT AGTCATGGAA
16501 AACCATTCCA ATCCTTATTT TCAATATATT TAAAAAGACA GAATTGACCC
16551 TGTTAACAGG CCTACCCTAA GAATCTTAAG AGCTTGCTTC CAGTTTGTCC
16601 TTGCTGCCTT CTGTATGCCT TGATTTCCCT GGAATTTAAG AGAAAGGATG
16651 TTATGGTACA GACCAAGTAG ATGACATAAA TGAACACCAC CTAAATCAG
16701 AGTTTTAAAA ATAGGCCCTG AACTGAAGCA AGAGGTAAAC TAGGGAAGCC
16751 TCAGGAGAAC TGAGACTTCT CCAGAGAGAA GTATCTGGGA TTAACTTCT
16801 TTCTAATGAG GCTTGGTTTT CCATGAACCT TTCCTTTAAA CCAAGGGGGG
16851 TATTGCTCAT CTTTCTGTTG AGCCCCATTT GTCATAATTG TAAAATGGGT
16901 GGTTACATCC TTCTGGTGAT CTAGGAGCCC TATTTTCGTC CTAGCATACA
16951 GCATTTTTCT AAAATTTGCT GTTAGCTTTC ATGATTCTTA CCCTAACTAT
17001 TCTTTTTCTA AAAAACATTT GTTTCAGCTT TACCACTCTG ATGAATTCAG
17051 AGCTTATGAC TGGGGAAATG ACGCTGATAA TATGAAACAT TACAATCAGG
17101 TGAGCTATTT ACAGTAACCC CAGCATGCTG ATTTTGATAA ATTATAATAA
17151 AAAATTATTT GAGGGTGGAA AGACTCCTAC CTGTCATTTG GTGGCATTTA
17201 TACTGATAGA ACTTTTTTTT AAAAAAATTT TAATTTTAAT TTTAATTTAT
17251 TTCAGAAAAT TTATAAATTA AAGAAGCATA TACAAAGAAA CTTACATCAT
17301 GTGTAATCCT TCCATCCAGA GATAACTAGA TGTACTAACA TTTTGGTGTA
17351 TTTATTCCAA TTTTCTCAGT ATTATATTGC TTTTAGACAA CTTTTAATCT
17401 TTCTATTTTA CTTAAGCTAT AGTAAGAGAT AACTAATATA ACTGAGGGAT
17451 TTTTAAATGC ATTTTTAATG GCTACATAAT AGAAATTATT TCATAAAAAT
17501 CTTTACAGCA TAAATGAATA TACACTTTTT AATACCAACA GAAAAATTAG
17551 AATTCCATAT GAAAGTTGAA TAAGTATTAC CCAACATTGA AGACTTGGGT
17601 CGTAAGGCAT CTTTCTCCAT ATAGCTTTAT GACATAAAAA TCTGTAGCCT
17651 TGTTTAGCAC CGTACTTTTA ATTAATCCTG TCACCATTTT TCTGTTCTCA
17701 TAGCCAGGGG CTTGGCTTAT AAGTATGAAC TAAGCAAACCT AAATTTAAAT
17751 GTTTTAAAGTA TTTTCCAGG CTATCATATT TTAAGCTATT TACTGGTGCA
17801 ACTATAGATT ATTAATAAGT TGTTTCTGAG GATCAAAAACA ATCAGACTAA
17851 TCAATTTCTC AATAATGAAT TGGCCTGTTA GAGGAATAAT TCTACTAATC
17901 CTTAAAACCA CTACAAGAGA TAGACCATGT ATATTTTATT TATTTTAA
17951 AATAAGTTTA AGATGTGATT TACATACAAG AACATTACTA ATTTTGTGTG
18001 TCCCATTTAA TAAGTTTGA CAAATATATT TATTTGTGTA ACCACACCAC
18051 AATCTAAATA TAGGACGTTT ATATCACCAC TAAAAGTTTT TTTCTGCTC
18101 CTGAGACTAT TTATAGACAC AAATGCGTGT ATTTGCAAAT GCTTAGAAAA
18151 GGTCTAGAAA AAAAAACAGT AAATGTTAAA GTGGTTATCT TCAGAGAGAA
18201 GAAAGAAGAA AAGAAGTGA TGGACATGAA ACAGTAAAGG ACCCTCATTT
18251 TGGACTTTAC ATATGTCTGT TTTCTTCCAT TATTTTGAAT AAACATGCTA
18301 TATTTATAAA TTATTTACAT TTACAAGAAA ATGAAACAAA ATCAACACGC
18351 ACATTCAAGA TCATTATGGT CAAGTACTAA AGTATGTGAG AGTGTTAATG

```

FIG. 3H

18401 TCCTTAGAAT TTGGCCACAG TTAGCTGGTC CTACTCTGCT CCAAGCCGGT  
18451 CCTATTTTGT GAATTAATCT CATTTGATGC CAATTTTAT TACATTCTCT  
18501 CCAAAAACT AGTCTCAACA GTTTGCTCTC TCCTCAAGTT CACAGCATT  
18551 TCTCTGCTAT ATCTATATTT TATTGAGTAT AAGAGAATTA ACCCATGTAA  
18601 GCTCCATGAG GGTAGGGATT TCTCATCGTT TTGTTACCA GTGTTTTCTC  
18651 ATCTTGAAGA GTACATGACA ATTACTGGGC TCCAGTATC TATGTGTTGC  
18701 ATTAATGAAA TTTCTTAAT TTAATCTACC TCAAATGTC TCTATCTTCT  
18751 TGATTCTCTC CTTCTTTTCT CTATCAGAAA ATGATGGTCC TCTTATTTTC  
18801 CAAGTTATTC CGGTCCTGTG CCCTTGATCC CATCTCTTCT CACTCCCCT  
18851 TCCTTCCTGC CTCCATTCTC CTGTCCCTTA TGAAAAACAA GCAAGACCAT  
18901 CAATTCTATC AAGTTATCAT TATGTCACTC TGTTCTTATC AACATATTTT  
18951 TAGTATTGAA GAGGGCTTCT TCTACTTACT CCTGAACCTT GTACAATGTA  
19001 GTTTAGGTCT TCATCTTTT ATCATAGCTA CCTTATTTAA AGTCACCCAT  
19051 GGCTTTTAAT TGCCAAATTC AATGGCCTAT CTTACCTTT TGAAATGTGT  
19101 TATGTTCTGT ACCACAGTCT CCTTGAACT CAGTCCCCTG ACTTGACTT  
19151 CCATAACACA ATGATTTCTG ATTTTCCTTC TGTTTGTGAT TGTTCTTTT  
19201 GTCCCAGGCA CTGGCTACTC CACCTTCCAC CTCTCTGAAA TCATTAGCAT  
19251 TCCCCAAGGA TTCTTCAAAA CTCTCTTTCT TCCTTGGAGA AGTCAGCATA  
19301 GCTTTAATTT GGACCATTTC TATGGCTTAT CTAGATTTT TCAGGACTTG  
19351 CTTCAACCT ATTCTTTCTG TAGGTGATT CATTAACTGT TGCCCATATG  
19401 GTAGTCCGAA GACAGACCTC CGAGAAATGA CCCTTGTCTC CAAAACCTCC  
19451 GCAATATGTC CAAATTTCT AGCCTGACAT TCAGACTTTG ATTATCTGCC  
19501 TCCAAGTTTA TATCCTATCA TATTCCTTTA TATATTCTGT TCTCCAGGTA  
19551 CACTGGGAAG CTTGCCATTC CTGATCATAG CCTACAACT CTTCTGCCT  
19601 CCCACTCACC CTCATCTCTG CTGTCAAAT GCAACCTTCC CTCAAGAGTC  
19651 ATTTACAGG ACCCTCTTT CTATGAAGCC CTCAGGTGGA AATAATTTTT  
19701 TGCTTTTTT TCCATTTTAT TTTTGGAGTG TTTATGGCAT TTAACATACC  
19751 TACTTTTGT TACAAATATT TGCCTTGCTC CCTCTTTTGC AAATTTCTTA  
19801 AAGGTAGAGA CCATTGTATG TTTTCTTCAT ATGTTGCTGG TGCCTAACAG  
19851 AACTATGGCC ATTGTCCACA TTCATTTAGC AGCCTTTGTA GTTATTGCTT  
19901 TGAGGAGCTT CCTCTCATGA ATGCCCTTGC TTTCTCTCCC ACAGAGTCAT  
19951 CCCCCTATAT ATGACCTGAC TGCCATGAAA GTGCCTACTG CTATTTGGGC  
20001 TGGTGGACAT GATGTCCTCG TAACACCCA GGATGTGGCC AGGATACTCC  
20051 CTCAAATCAA GAGTCTTCAT TACTTTAAGC TATTGCCAGA TTGGAACCAC  
20101 TTTGATTTTG TCTGGGGCCT CGATGCCCT CAACGGATGT ACAGTGAAAT  
20151 CATAGCTTTA ATGAAGGCAT ATTCCTAAAT GCAATGCATT TACTTTTCAA  
20201 TTAAAAGTTG CTTCCAAGCC CATAAGGGAC TTTAGAAAAA ATGGTAACCA  
20251 ACAATGAGGT TGTCCCCCAG CACCCTGGGG GAGATGCACA GTGGAGTCTG  
20301 TTTTCCAAGT CAATTGTGTT AGTGTTATTT ATGTTTAGAG ACATCTTTGC  
20351 ATGGGACCAT CTACAGGTCC TTATAAACAA TGAGGTAGAT TAGGCAAAAA  
20401 GATAAACAAAG TTGCTACTCT ATCTGGCATT TAAGTCTAAT TAAATTGTAA  
20451 TTTTtagggc ATACCATGAA GTATAGAAAT GTCTGAAGCT TCAAAGGAAC  
20501 AGTGAAATTC CTTTAAGGTC CTATATGGAA ACCTCTGTTG TCATTTTATT  
20551 TATATGGATT GCTATGGCAA TGGACAGAGT GTGGGATTAG GAGGAGGGCC  
20601 TGTAACCTCT TTATAAAAGT TTCTTAGCTA TCCTGAAGAT GTATAGACAT  
20651 TTTTACTTTT TTAGGTATTT TCAACATCAG AAATTCAAAA AAGTCCCCAA

FIG. 3I

```

20701 AGATTCTTCC AGAGAAGCCC TCTTTTCTTA CAATCTTATC CCTGGCTATC
20751 TGC GTAAACG GAATCTTGAA CCCATAATAG GATACATGTA TAAAATCTTC
20801 CTTATTAAAG CAGAAATAAA TTGTACAGCA TCAATATCAT TTTATAATCA
20851 TAGGGAGGCT TCTTTGTTTA GCATGTAATG CCCCCTTTAC AGGCTTTTTG
20901 TTCTTTGAGG GGT TTGAACA TTCCATGAAA AACTGACAGA TAGGAAACTG
20951 ACAATAAAAG ATTGAGCTAA AGATGGAAGC AGAAAGTACT AGGCTAGATA
21001 GTCTCTAAAC ATTAAGTATT TTCTTCCTCC ATCTTAAAAG CAATGAGAAG
21051 CCACCAAAAT ATTTTACCTA ATGGAAACCT GATTGCCGCA TTTTGTAAAC
21101 CACCACTTTG GCTGCTACAT AGAGAATGGA TTAGAAGATG CCAACAAAAG
21151 ATTCTGAGCA AGTCTGTAAG TCTGATCAAG TGTTCTGATG CAGGCTGATA
21201 TCCTTCTGTG CTAAGAGAGA TGATCCTTGG AAAATCCAGA GCCAGCTCCA
21251 TAATACTTTC CTGCTCTGCT GGCAAATCCA CAAGCTGCTG GCCCCTGGAG
21301 CCATTCTTCT CTCAAACTA GCATTCATCA ATTTAATGTA TACGTATTGA
21351 TGGGGAATAA TGGTCACTAT GAAAACCATG TGATAATATG GAAAAATACC
21401 CATGATATAA TGTTATGTGA AGAGAAGAAA ATGAAACTGG TAGAACTATG
21451 TGATTGCAAA TATATACAAA TATTAAAACA ATTATATGAC TTTATAAAAT
21501 ATTTGTATAT AATGAAAACCT GAAGCAATAT AAAAAATAAA ATTAGTTGTG
21551 TCAGGGTAGT AACATGATGA GTGATTAATA GTTTTAAATT TTTAATATAG
21601 TAATGACATA ATGTTACAAC TTGTCCAAAT CTCACAAACA TAATATTCAG
21651 TAAAGGAAGA TAAACATAAA AGAATACATA TTTTATTATA CATTTTTATG
21701 TAGGCTAATT GATGGTTCTG AAAGCCTTAA AAAGCTTACT TTTAGGAGGA
21751 GAATCATGCC TTGGAGGACT CTAGGTCCA GAAAAATGTC CTAATACTAG
21801 AGCTAGGTGC AGTCAGATTA ATTATAATAC ATTTTATTAT TTTGTCTGGA
21851 ATACCAAGAT GACTTCCAAG CAGGAATGGA GTCTAGCAAC ACTTTACTGA
21901 TGGGGAACCT GGCCACAGAC TTGTAATACA AATTTTGGGA TATGTTGACA
21951 ATGTTTCTCC TTATTTTCT TACTTATACA AAGCAAGAAA TTTGGCTCAC
22001 AACCTTGAAA CAGACTTACC AGGTTCTCC AGTTTCCCA GCCTCAATAT
22051 CTCATTGCTA TTTTAA

```

(SEQ ID NO: 3)

#### SNPs:

DNA		
Position	Major	Minor
165	G	A
226	A	G
231	T	C
359	A	-
544	G	T
598	C	T
1621	A	G
2330	C	T
2498	A	G
2791	T	C
2877	T	C

FIG. 3J

2879	T	C
2912	A	G
3076	G	T
3745	C	G
3752	T	-
3762	-	C T
3833	A	G
4399	T	C
4945	A	G
5056	A	G
5280	T	A
5790	A	G
5901	C	T
6457	C	T
6632	T	A
6763	A	G
6955	-	T C
7017	T	G
7151	G	T
7308	C	G
7321	T	C
7542	C	T
8597	T	C
8803	C	T
9016	G	A
9967	T	C
10008	C	T
10363	G	A
10684	T	C
11177	G	T
12345	T	C
12349	C	T
13115	C	T
13354	T	A
13373	C	G
14677	C	G
14734	G	A
14747	A	G
14808	-	A
15086	-	A G
15414	A	G
15722	T	C
15861	T	C
16264	A	T
16314	G	A
16877	A	G

FIG. 3K



16966	T	G
17147	A	G
17219	T	C
18628	A	G
18655	T	G
18984	G	T
19407	C	T
19531	T	C
19911	C	T
20199	A	G
20243	G	A
20640	T	C
21156	G	C
21163	A	T
21425	G	A

Context:

DNA

Position

165 TTATGGCCTAACCTTTTAACTTTGAGTTATTTTCAAGAGAAAATTTGAAAAAGCAGCCT  
TTGAGGAGAAAGAAGCAATCCAACAAACAAAAGATAACCACACTGTAATAGGAAATGTG  
TTTTGAATAGGACATTGGAAGAAAAATAATAATCATTTTTTACAG  
[G,A]  
TAGATCCCAAAGTCAAGGATCTATGTTCAACCATGTGTGTTCCACCATCTTCACAATTGA  
ATGAGTAACCATCATTAAGCAGTTAGCTTAGGCCGTAATATGATTCTTGGACTGAGATTT  
CAAAAATACCACAGGCCTTCTGAAAGGTTACCCCTTTCTAGCTCCACTATCATCTAATTT  
TATTAATAAAAAAAAAAAGGAAAAATTTGAGCTTCTAGAGAGTAGGGGCTACCATTTTG  
TATCCACAGGGCCAAGGAACAAGTTTTAATGTATTCATTTAAATTAATTTTCAGTATGAG

226 TTATGGCCTAACCTTTTAACTTTGAGTTATTTTCAAGAGAAAATTTGAAAAAGCAGCCT  
TTGAGGAGAAAGAAGCAATCCAACAAACAAAAGATAACCACACTGTAATAGGAAATGTG  
TTTTGAATAGGACATTGGAAGAAAAATAATAATCATTTTTTACAGGTAGATCCCAAAGTCA  
AGGATCTATGTTCAACCATGTGTGTTCCACCATCTTCACAATTGA  
[A,G]  
TGAGTAACCATCATTAAGCAGTTAGCTTAGGCCGTAATATGATTCTTGGACTGAGATTT  
AAAAATACCACAGGCCTTCTGAAAGGTTACCCCTTTCTAGCTCCACTATCATCTAATTTT  
ATTAATAAAAAAAAAAAGGAAAAATTTGAGCTTCTAGAGAGTAGGGGCTACCATTTTG  
ATCCACAGGGCCAAGGAACAAGTTTTAATGTATTCATTTAAATTAATTTTCAGTATGAGT  
ATTGAAATATATAATAGAAATATTGTAACATTATATATTTTCTATATACTTTTATTATAT

231 TTATGGCCTAACCTTTTAACTTTGAGTTATTTTCAAGAGAAAATTTGAAAAAGCAGCCT  
TTGAGGAGAAAGAAGCAATCCAACAAACAAAAGATAACCACACTGTAATAGGAAATGTG  
TTTTGAATAGGACATTGGAAGAAAAATAATAATCATTTTTTACAGGTAGATCCCAAAGTCA  
AGGATCTATGTTCAACCATGTGTGTTCCACCATCTTCACAATTGAATGAG  
[T,C]

FIG. 3L

AACCATCATTAAGCAGTTAGCTTAGGCCGTAATATGATTCTTGGACTGAGATTTCAAAAA  
TACCACAGGCCTTCTGAAAGGTTACCCCTTTCTAGCTCCACTATCATCTAATTTTATTAA  
AAAAAAAAAAAAAGGAAAAATTTGAGCTTCTAGAGAGTAGGGGCTACCATTTTGTATCCC  
ACAGGGCCAAGGAACAAGTTTAAATGTATTCAATTTAAATTAATTTTCAGTATGAGTATTGA  
AATATATAATAGAAATATTGTAACATTATATATTTTCTATATACTTTTATTATATAGAAA

359 CTTTGAGGAGAAAGAAGCAATCCAACAAACAAAAAGATAAACCACACTGTAATAGGAAATG  
TGTTTTGAATAGGACATTGGAAGAAAAATAATAATCATTTTTACAGGTAGATCCCAAAGT  
CAAGGATCTATGTTCAACCATGTGTGTTCCACCATCTTCACAATTGAATGAGTAACCATC  
ATTAAGCAGTTAGCTTAGGCCGTAATATGATTCTTGGACTGAGATTTCAAAAATACCACA  
GGCCTTCTGAAAGGTTACCCCTTTCTAGCTCCACTATCATCTAATTTTATTAAAAAAA  
[A, -]  
AAAAAGGAAAAATTTGAGCTTCTAGAGAGTAGGGGCTACCATTTTGTATCCACAGGGCC  
AAGGAACAAGTTTTAATGTATTCAATTTAAATTAATTTTCAGTATGAGTATTGAAATATATA  
ATAGAAATATTGTAACATTATATATTTTCTATATACTTTTATTATATAGAAAATATATAT  
TACAGAATATATTATTAATATTGTAGAACAATATATAATACAGAAAAATATATAATACT  
CAGTAATATATTAAATACTTATTAAAATAGCAAGCTTATATAGGAAGAGTGATGGAGCAT

544 GCAGTTAGCTTAGGCCGTAATATGATTCTTGGACTGAGATTTCAAAAATACCACAGGCCT  
TCTGAAAGGTTACCCCTTTCTAGCTCCACTATCATCTAATTTTATTAAAAAAA  
AGGAAAAATTTGAGCTTCTAGAGAGTAGGGGCTACCATTTTGTATCCACAGGGCCAAGG  
AACAAGTTTTAATGTATTCAATTTAAATTAATTTTCAGTATGAGTATTGAAATATATAATAG  
AAATATTGTAACATTATATATTTTCTATATACTTTTATTATATAGAAAATATATATTACA  
[G, T]  
AATATATTATTAATATTGTAGAACAATATATAATACAGAAAAATATATAATACTCAGTA  
ATATATTAATACTTATTAAAAATAGCAAGCTTATATAGGAAGAGTGATGGAGCATTGTGA  
GAAAGTTTCAGCTTTATTTCTTTGACATTACTTTGTTTCTGCACAAACAAAAGAATTACA  
GGAATTGTCCAGATTATTCAAATAACTCGAAGTTGAGGAGGGAATATAAGTCAATGATGT  
AGAAACTCTTTAAGATTTGAGCTAGCCTACAATCTGTAAAGATCTGTGAAATTGAACTA

598 AGGCCTTCTGAAAGGTTACCCCTTTCTAGCTCCACTATCATCTAATTTTATTAAAAAAA  
AAAAAAGGAAAAATTTGAGCTTCTAGAGAGTAGGGGCTACCATTTTGTATCCACAGGG  
CCAAGGAACAAGTTTTAATGTATTCAATTTAAATTAATTTTCAGTATGAGTATTGAAATATA  
TAATAGAAATATTGTAACATTATATATTTTCTATATACTTTTATTATATAGAAAATATAT  
ATTACAGAATATATTATTAATATTGTAGAACAATATATAATACAGAAAAATATATAATA  
[C, T]  
TCAGTAATATATTAATACTTATTAAAAATAGCAAGCTTATATAGGAAGAGTGATGGAGCA  
TTGTGAGAAAGTTTCAGCTTTATTTCTTTGACATTACTTTGTTTCTGCACAAACAAAAGA  
ATTACAGGAATTGTCCAGATTATTCAAATAACTCGAAGTTGAGGAGGGAATATAAGTCAA  
TGATGTAGAAACTCTTTAAGATTTGAGCTAGCCTACAATCTGTAAAGATCTGTGAAATT  
GAACTATATTTGTGCTATTTCCATATTAAGTCAAGGCAACAAATCAATATTAATAATAAT

1621 CGGCTTAAGCTCCACAGGCATACAAAGTGAAGCAGAAAACTGAGGCACGTGTGCCTCCAT  
TATCTGGTATCTCATGTGGGGCTTAGAGGTAAATTGTCGTTATTTGGCCTCCATTTCTGC  
CTTTAACCCTGGTGTAACAAAGGTTACTGTGCCAAAGTTGACAGCAACCCAAATCCCT  
TTGGCATGTGAATTAGTTTCCTCTGCCATACTGCTAGTTCCAAATTCCTTCTGGTTTCAG

FIG. 3M

GATTTAGGAGTCAGGGTTGCCTCATCTTCTCAAATGAGTTACAGTCACGCACATCCCTAC  
[A,G]  
CACTGCATGGTTGGCACTAGTTCCTTGATATATGTTACTCCGTTTGATCCTCATGAAGGA  
TCAAATGGGGAAGGGAGATACTATTGTCTCTGATTGTCCATTAAGATCTTGAGTATGTTCT  
TACTTCCCTGTTTGACACACTGGTTTGAAAATGTTGCTAAGTCTTCCCAACAATGACAGA  
TACTCAGTGGAACATGAAGGATTCCGTCAAACCTGGTTATTTTGCATCATGTAGACCACT  
ATTTCCCAACCTGCAAGTGCATCATGGCCTTTGGTGTGTCAGGGACACGCCTTGGGTGTG

2330 AAAAGTTCAGAAGTTCCTCATCAATAAGGAGTCCTTGTGAGCAGGTGAAGCTCATCTAAC  
TAGGTAAGATGAAGATCTATCATAACCAGGAGGCAGGTTGGAAGGTGCCAGTTGCACTGG  
CAGTCAGGTGCAAGAGCTCTGCAGTGAGGCTGCCTGAGTGTCCATCCTAGATCTCTCACC  
TCTTGGCTCTGTGACCTTGAGCAGGTCTTAAATCTCTCTAAGCCTTTGTTTTTTAATTG  
ATAAAATGAGGATAATAATAGTACCAAATTAGGGAGATTTTCAGAGCTTAAATAACATA  
[C,T]  
GTGAACTATTTAGAGTAATGCCTGCCATAAGGGGACTCAGTAGCTTATTATTAGTTTCAT  
ACAATTTGAAAAGTTTCATAATATTTGCAGATATAAGATGATCTTCAACCAGATAGCTAA  
TGTATGCAAAGCTATTTAGCTTCAGAAGTAACTCTGCATTTCTAGAAGTTAAATATTAC  
TTTGTATAGTGAATTATCTGTAATATTTATCTCTTGCTCACTTTTATAAGAAAAATAGT  
GAAAGCATTTATTAAGAACTTACACTGCACTAAATGTTATATATGACTTAATCCTCACTA

2498 AGATCTCTCACCTCTTGGCTCTGTGACCTTGAGCAGGTCTTAAATCTCTCTAAGCCTTTG  
TTTTTTAATTGATAAAATGAGGATAATAATAGTACCAAATTAGGGAGATTTTCAGAGC  
TTAAATAACATACGTGAACTATTTAGAGTAATGCCTGCCATAAGGGGACTCAGTAGCTTA  
TTATTAGTTTCATACAATTTGAAAAGTTTCATAATATTTGCAGATATAAGATGATCTTCA  
ACCAGATAGCTAATGTATGCAAAGCTATTTAGCTTCAGAAGTAACTCTGCATTTCTAGA  
[A,G]  
GTTAAATATTACTTTGTTATAGTGAATTATCTGTAATATTTATCTCTTGCTCACTTTTAT  
AAGAAAAATAGTGAAAGCATTTATTAAGAACTTACACTGCACTAAATGTTATATATGACT  
TAATCCTCACTATAACCCTATGAGATAGGTTACATTATTGTCCTAATTTTACTAACAAGG  
AAACCAAGAGACAAAGCTACTAAACACTTGCCTGAGGTAGACATCTTCTCTGTGGTG  
AGGCTGGATTTCAAATTTAGACCATTTGACTGTAGCACTTATATGATGAGCATGCTGTTT

2791 TTCTAGAAGTTAAATATTACTTTGTTATAGTGAATTATCTGTAATATTTATCTCTTGCTC  
ACTTTTATAAGAAAAATAGTGAAAGCATTTATTAAGAACTTACACTGCACTAAATGTTAT  
ATATGACTTAATCCTCACTATAACCCTATGAGATAGGTTACATTATTGTCCTAATTTTAC  
TAACAAGGAAACCAAGAGACAAAGCTACTAAACACTTGCCTGAGGTAGACATCTTCTT  
CTGTGGTGAGGCTGGATTTCAAATTTAGACCATTTGACTGTAGCACTTATATGATGAGCA  
[T,C]  
GCTGTTTAGTGTTATAGTGTTGGTCTACCTTTGAATAGACATACTTTTAAACCATGGCAA  
GGAAGTGAGACTGCACATTGAAATATGTAAATTTGCCTTTGGGTGCCACGTGAGAAATA  
GTCACATCACTAGAACTAATCATAAGCTTTTGTGTTTGGTTAAAGTTTTATTGATCCAT  
TTTTCTTGTTTACTTTGTGGGATACTGGGCTTAACTAGGGGATACCTCCACTTTTTACTT  
GGCCATGGTATGAAAACCTGTCCTCTGAATCTTTAGATATTTTGGCAAATGTAGGCAAA

2877 ATTTATTAAGAACTTACACTGCACTAAATGTTATATATGACTTAATCCTCACTATAACCC  
TATGAGATAGGTTACATTATTGTCCTAATTTTACTAACAAGGAAACCAAGAGACAAAGCT

FIG. 3N

ACTAAACACTTGCCTGAGGTTAGACATCTTCTTCTGTGGTGAGGCTGGATTTCAAATTT  
AGACCATTTGACTGTAGCACTTATATGATGAGCATGCTGTTTAGTGTTATAGTGTTGGTC  
TACCTTTGAATAGACATACTTTTAAACCATGGCAAGGAAGTGAGACTGCACATTGAAATA  
[T,C]  
GTAAAATTTGCCTTTGGGTGCCACGTGAGAAATAGTCACATCACTAGAACTAATCATAA  
GCTTTTGTGTTGGTTAAAGTTTTATTGATCCATTTTCTTGTTTACTTTGTGGGATACT  
GGGCTTAACTAGGGGATACCTCCACTTTTTACTTGGCCATGGTATGAAAACCTGTCCTCT  
GAATCTTTAGATATTTTGGCAAATTGTAGGCAAACAAAGACTTAAAGCAATTCAACCTTG  
ATAAAATAAGACCAAAAATGCCTCCATACTTGATTAAATTTATTTTCAATTTAGGAACTG

2879 TTATTAAGAACTTACACTGCACTAAATGTTATATATGACTTAATCCTCACTATAACCCTA  
TGAGATAGGTTACATTATTGTCCTAATTTTACTAACAAGGAAACCAAGAGACAAAGCTAC  
TAAACACTTGCCTGAGGTTAGACATCTTCTTCTGTGGTGAGGCTGGATTTCAAATTTAG  
ACCATTTGACTGTAGCACTTATATGATGAGCATGCTGTTTAGTGTTATAGTGTTGGTCTA  
CCTTTGAATAGACATACTTTTAAACCATGGCAAGGAAGTGAGACTGCACATTGAAATATG  
[T,C]  
AAAATTTGCCTTTGGGTGCCACGTGAGAAATAGTCACATCACTAGAACTAATCATAAGC  
TTTTGTGTTTGGTTAAAGTTTTATTGATCCATTTTCTTGTTTACTTTGTGGGATACTGG  
GCTTAACTAGGGGATACCTCCACTTTTTACTTGGCCATGGTATGAAAACCTGTCCTCTGA  
ATCTTTAGATATTTTGGCAAATTGTAGGCAAACAAAGACTTAAAGCAATTCAACCTTGAT  
TAAAATAAGACCAAAAATGCCTCCATACTTGATTAAATTTATTTTCAATTTAGGAACTGGA

2912 TATGACTTAATCCTCACTATAACCCTATGAGATAGGTTACATTATTGTCCTAATTTTACT  
AACAAGGAAACCAAGAGACAAAGCTACTAAACACTTGCCTGAGGTTAGACATCTTCTTC  
TGTGGTGAGGCTGGATTTCAAATTTAGACCATTTGACTGTAGCACTTATATGATGAGCAT  
GCTGTTTAGTGTTATAGTGTTGGTCTACCTTTGAATAGACATACTTTTAAACCATGGCAA  
GGAAGTGAGACTGCACATTGAAATATGTAATTTGCCTTTGGGTGCCACGTGAGAAATA  
[A,G]  
TCACATCACTAGAACTAATCATAAGCTTTTGTGTTTGGTTAAAGTTTTATTGATCCATT  
TTTCTTGTTTACTTTGTGGGATACTGGGCTTAACTAGGGGATACCTCCACTTTTTACTTG  
GCCATGGTATGAAAACCTGTCCTCTGAATCTTTAGATATTTTGGCAAATTGTAGGCAAAC  
AAAGACTTAAAGCAATTCAACCTTGATTAAAATAAGACCAAAAATGCCTCCATACTTGAT  
TAAATTTATTTCAATTTAGGAACTGGATTATAATCAAGACAACCTTCTACATGAAAAAATA

3076 CTTATATGATGAGCATGCTGTTTAGTGTTATAGTGTTGGTCTACCTTTGAATAGACATAC  
TTTTAAACCATGGCAAGGAAGTGAGACTGCACATTGAAATATGTAATTTGCCTTTGGG  
TGCCACGTGAGAAATAGTCACATCACTAGAACTAATCATAAGCTTTTGTGTTTGGTTAA  
AGTTTTATTGATCCATTTTCTTGTTTACTTTGTGGGATACTGGGCTTAACTAGGGGATA  
CCTCCACTTTTTACTTGGCCATGGTATGAAAACCTGTCCTCTGAATCTTTAGATATTTTG  
[G,T]  
CAAATTGTAGGCAAACAAAGACTTAAAGCAATTCAACCTTGATTAAAATAAGACCAAAAA  
TGCCTCCATACTTGATTAAATTTATTTTCAATTTAGGAACTGGATTATAATCAAGACAAC  
TCTACATGAAAAAATAGATTAATAGTGCTCCAAGTTAGTTCACTGTATTTATTCCTTTTT  
ATACATTATCTGCCTTCGGTGTTATTCAAGTTTTTCAATTAATCAATAATTTCACTAAT  
CATTTTATTTCAATTAATCAACATTGATAGTTAAATTAATCTGTGAATATTAATGTTTT

FIG. 30

- 3745 TGGTGGATTCTTGATTGGAAAATGAAGTGAATCCTGAGGTGTGGATGAATACTGTAAG  
TCATGGAAAACGTGAAGAACATCAAATAAAGCAGGACTAATGGAGTATGAGGTTACGAA  
AGGTCCTGTTGTAAACAGAAAATCTCTGATAAAACAGATAAAATGTAGATGGTTTTTAACC  
TCTGCAAGAGTCAAGCTAGTTAGATCTTTGTCTGAAAAACAAATACTGTCCGGTAATGAA  
AACCAAATTGTGCTATTGTGCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT  
[C,G]  
TATCTATCTATCTATTTATCTATCTATCTATAGATAGAACCTCCTCTTTTGAATTTATGT  
TTTAAGAATATCAAGCTATTTGTTGATATACATGATTGCCTTCTATTGATCTATAGTTCT  
ATTACTTTTAAAGCAAGAGGGGTCTCAAAGACAATTGACTTGATAATATAGCTTTGTCA  
GAAAGAATGGGTCAATGCTAAATTTTCCCCCAACCCCCCAAAATATTAGCCAATAGTAGA  
TATTTTTTAAATTTCTACTTATTTTGTATTAAGACTTTATTTATTAATTTTACAGTTACC
- 3752 TTCCTTGATTTGGAAAATGAAGTGAATCCTGAGGTGTGGATGAATACTGTAAGTCATGGA  
AAACTGTGAAGAACATCAAATAAAGCAGGACTAATGGAGTATGAGGTTACGAAAGGTCCT  
GTTGTAAACAGAAAATCTCTGATAAAACAGATAAAATGTAGATGGTTTTTAACCTCTGCAA  
GAGTCAAGCTAGTTAGATCTTTGTCTGAAAAACAAATACTGTCCGGTAATGAAAACCAAA  
TTGTGCTATTGTGCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA  
[T,-]  
CTATCTATTTATCTATCTATCTATAGATAGAACCTCCTCTTTTGAATTTATGTTTTAAGA  
ATATCAAGCTATTTGTTGATATACATGATTGCCTTCTATTGATCTATAGTTCTATTACTT  
TTAAAGCAAGAGGGGTCTCAAAGACAATTGACTTGATAATATAGCTTTGTGAGAAAGAA  
TGGGTCAATGCTAAATTTTCCCCCAACCCCCCAAAATATTAGCCAATAGTAGATATTTTT  
TAAATTTCTACTTATTTTGTATTAAGACTTTATTTATTAATTTTACAGTTACCTGGTGCT
- 3762 TGGAAAATGAAGTGAATCCTGAGGTGTGGATGAATACTGTAAGTCATGGAAAACGTGAA  
GAACATCAAATAAAGCAGGACTAATGGAGTATGAGGTTACGAAAGGTCCTGTTGTAAACAG  
AAAATCTCTGATAAAACAGATAAAATGTAGATGGTTTTTAACCTCTGCAAGAGTCAAGCT  
AGTTAGATCTTTGTCTGAAAAACAAATACTGTCCGGTAATGAAAACCAAAATTGTGCTATT  
GTGCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATT  
[-,C,T]  
ATCTATCTATCTATAGATAGAACCTCCTCTTTTGAATTTATGTTTTAAGAATATCAAGCT  
ATTTGTTGATATACATGATTGCCTTCTATTGATCTATAGTTCTATTACTTTTAAAGCAAG  
AGGGGTCTCAAAGACAATTGACTTGATAATATAGCTTTGTGAGAAAGAAATGGGTCAATG  
CTAAATTTTCCCCCAACCCCCCAAAATATTAGCCAATAGTAGATATTTTTTAAATTTCTA  
CTATTTTTGTATTAAGACTTTATTTATTAATTTTACAGTTACCTGGTGCTACAAATTTCA
- 3833 AAAGCAGGACTAATGGAGTATGAGGTTACGAAAGGTCCTGTTGTAAACAGAAAATCTCTGA  
TAAACAGATAAAATGTAGATGGTTTTTAACCTCTGCAAGAGTCAAGCTAGTTAGATCTT  
TGTCTGAAAAACAAATACTGTCCGGTAATGAAAACCAAAATTGTGCTATTGTGCTATCTAT  
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATTTATCTATCTAT  
CTATAGATAGAACCTCCTCTTTTGAATTTATGTTTTAAGAATATCAAGCTATTTGTTGAT  
[A,G]  
TACATGATTGCCTTCTATTGATCTATAGTTCTATTACTTTTAAAGCAAGAGGGGTCTCAA  
AAGACAATTGACTTGATAATATAGCTTTGTGAGAAAGAAATGGGTCAATGCTAAATTTTCC  
CCCAACCCCCCAAAATATTAGCCAATAGTAGATATTTTTTAAATTTCTACTTATTTTGTAT  
TTAAGACTTTATTTATTAATTTTACAGTTACCTGGTGCTACAAATTTTACAGATAATTCACC

FIG. 3P

CTAATAAGCACACAACAGATGGTTTGTTTTGATTCTTTTTATATCCTTTGGAGAAGTTC

4399 GTTTTGATTCTTTTTATATCCTTTGGAGAAGTTCCTAACGACTGTATTTTACTGGG  
CAGAGTGAATCATCATCTACAATGGCTACCCAGTGAAGAGTATGAAGTCACCACTGAA  
GATGGGTATATACTCCTTGTCAACAGAATTCCTTATGGGCGAACACATGCTAGGAGCACA  
GGTACAAGATATGTCTCTCCTGAAAAGGGGACTGCATTGACCTCCTGCTTCTCAGGAGGA  
ATTTAATGCTAGATATGCATCAACAGAGTTTATCAAAATTGGTTTGAATTATTGGATTAG  
[T,C]

CTTTAAATAGTTATCAGGGAGGCTCACTCTTTCCTGATAATTCTCTGAAGACAGACAGG  
AACCTAAAAATACAAACAGCAAGACTGATCTTGCTAACTGCAACCAGAGGTACTTGTTAG  
GGTGTAACAGAAAGGCAGAGCCTGCATTTTGTACCTCATTACTGATTTATCATGTGGA  
AAATTGCTTTGTCCCAGGAAAATGGATCCTCTCATTGTCAGAAGGAGATTTTCTAGGTTG  
TATGAAATTGACTCTGGGGCACCCAAGAAGAACCTCTCCTGCTCCCACTAAAATTAAGGG

4945 AATTGACTCTGGGGCACCCAAGAAGAACCTCTCCTGCTCCCACTAAAATTAAGGGGCCTC  
CCTCTGCAGGATAAAAAACAATCTAGTTAAATGACAACGCATTTCTGAAAAGTTTTCCAG  
GACTGAAAACCTTAACATCCACATACACTTTGATCTAAGGGACAGACGGTTCATAGAATG  
AAAGAGTATGGTGTCAATAAGGCTTGAATTCCTAGAATGAGGAGCCAGCCATGCCATAGCA  
GGGGAATGATACTCCTTAAAAGGGAAAATTTAACTACAAATCCTCTGAAGTAGAAATGAT  
[A,G]

AGAATAACCAAAATATCTGCAATGGTTCAATAGCAAATAATTTATTGGCAGCTGCTTACC  
GTGTTCATTTTGCATCTTTTTCCCACCACACATATTAAGGAGCAGCTGAAGTCATGTTT  
GACATTCTCTCCCTCTTTTATCTCCAGTTTCAGAATGAAAAATGAGAGTGAGATATGAGT  
AGTTTTACTAGTTAAATATGAAACACCCAGTTAAATTTGAAGGTCAGATAAACAACAA  
TAATTTTGTATAAGTCTCATTTTAAGATAATACTAAAAAGTCATTATTTATTCATTATTA

5056 GTTTTCCAGGACTGAAAACCTTAACATCCACATACACTTTGATCTAAGGGACAGACGGTT  
CATAGAATGAAAGAGTATGGTGTCAATAAGGCTTGAATTCCTAGAATGAGGAGCCAGCCAT  
GCCATAGCAGGGGAATGATACTCCTTAAAAGGGAAAATTTAACTACAAATCCTCTGAAGT  
AGAAATGATAAGAATAACCAAAATATCTGCAATGGTTCAATAGCAAATAATTTATTGGCA  
GCTGCTTACCGTGTTCATTTTGCATCTTTTTCCCACCACACATATTAAGGAGCAGCTGA  
[A,G]

GTCATGTTTGACATTCTCTCCCTCTTTTATCTCCAGTTTCAGAATGAAAAATGAGAGTGA  
GATATGAGTAGTTTTACTAGTTAAAATATGAAACACCCAGTTAAATTTGAAGGTCAGATA  
AACAACAAATAATTTTGTATAAGTCTCATTTTAAGATAATACTAAAAAGTCATTATTTAT  
TCACTATTATCACTATTTATAAAATTTTGTAGAGCATCCTGGATCTTTTTGCTTACTTTT  
GTTTTATTTTTTGCTAAATCTGGCAATCCCAGGCACATGTGTGAAGGAGCTGTGAAATA

5280 AAATAATTTATTGGCAGCTGCTTACCGTGTTCATTTTGCATCTTTTTCCCACCACACAT  
ATTAAGGAGCAGCTGAAGTCATGTTTGACATTCTCTCCCTCTTTTATCTCCAGTTTCAGA  
ATGAAAAATGAGAGTGAGATATGAGTAGTTTTACTAGTTAAAATATGAAACACCCAGTTA  
AATTTGAAGGTCAGATAAACAACAAATAATTTTGTATAAGTCTCATTTTAAGATAATACT  
AAAAAGTCATTATTTATTCATTATTATCACTATTTATAAAATTTTGTAGAGCATCCTGGA  
[T,A]

CTTTTTGCTTACTTTTGTTTTATTTTTTGCTAAATCTGGCAATCCCAGGCACATGTGTG  
AAGGAGCTGTGAAATATAAAAGGAGAAAACCTTTATGGGAAAGATTTGGCTTAAGGAGAG

ATAATTTTGGAAAGATTTAGAATTAAGATCATTATTAGATGTAATGTTCTAAATACTT  
TATATCAGTTAAACTTCTCATCAACAATATGAGATGGGTACCACTAATAGTCACCATTTC  
ACAAATGATGAAATTAAGGCACAACCGTTATGTTAAGAGGCCTAAAGTCCACAAATAGC

5790 TGAGATGGGTACCACTAATAGTCACCATTTCACAAATGATGAAATTAAGGCACAACCGGT  
TATGTTAAGAGGCCTAAAGTCCACAAATAGCAAGCTGACAGACCAGAATTTAAGCCCAGG  
CATGCTGGCTCCAGAGCCTGTGCTCTTAGTCATTAAATTATAGTGCCTTACTTGACCTTC  
CACCTGGTTACTTTGGATCTCCCTGAATGCTCTCTCTCCCTCAGAAATACTGGAAGTTG  
GCAGAGGGACACTGAGCTGAGCATATTATTGTAGTTTTTAAATGCTCTCCACTGGACAGA  
[A,G]  
GATGGGGGATTTGAATAGAAATTTGGTGAGGAACATAATCAGTGTCCATTTACACTCACCT  
CCTCTTCCCTCCCTGGAAGAGCTATAGGACTTGAGTAAGCATGATAAATTTCTGTCTTTG  
TAAACCACACCCAGGAAATTTGTATATACAAATACATAGAGCACAGTAGTTATCAGGACA  
GACTTTGACATAAAAAGAACTGGGTTTGAGTCCCTGCTCTGGCCTTCTTATCTGGGTGGC  
CCTCTGGGAAAGTTACTTAACTACATAAAGTTTTGTTTCCATATCTACAAAATGAGGTTT

5901 AAGCCCAGGCATGCTGGCTCCAGAGCCTGTGCTCTTAGTCATTAAATTATAGTGCCTTAC  
TTGACCTTCCACCCTGGTTACTTTGGATCTCCCTGAATGCTCTCTCTCCCTCAGAAATAC  
TGGAAGTTGGCAGAGGGACACTGAGCTGAGCATATTATTGTAGTTTTTAAATGCTCTCCA  
CTGGACAGAAGATGGGGGATTTGAATAGAAATTTGGTGAGGAACATAATCAGTGTCCATTT  
ACACTCACCTCCTCTTCCCTCCCTGGAAGAGCTATAGGACTTGAGTAAGCATGATAAATTT  
[C,T]  
GTGTCTTTGTAAACCACACCCAGGAAATTTGTATATACAAATACATAGAGCACAGTAGTT  
ATCAGGACAGACTTTGACATAAAAAGAACTGGGTTTGAGTCCCTGCTCTGGCCTTCTTAT  
CTGGGTGGCCCTCTGGGAAAGTTACTTAACTACATAAAGTTTTGTTTCCATATCTACAAA  
ATGAGGTTTCTCAAAATAGCAGCTAGTTTATAGAGTTGTTGCAAGAATTTAGTAAGCTAA  
TACATATAAATACGTCAACATAGCACCAGGTACAAAATATGTGCTCAAGAACTGAAGT

6457 CAACATAGCACCAGGTACAAAAATATGTGCTCAAGAAACTGAAGTTACCTGATTATAATG  
CTCTATACTATTGACAAGGGAAAAAGTGAAAACAGTTTTTGTGTTTACCATGTGTGTATGTG  
TGTGTGTCTGTGATGTTCCGACATGCTCTATTTAACATAAATTACTCTCACTCTTTCTC  
TCTCTCTCTTTCTCTTTCTCCCTCTCTCATCTTACCCTTTCCCCCACCAGGTCCCCGGCC  
AGTTGTGTATATGCAGCATGCCCTGTTTGCAGACAATGCCTACTGGCTTGAGAATTATGC  
[C,T]  
AATGGAAGCCTTGGAATTCCTTCTAGCAGATGCAGGTTATGATGTATGGATGGGAAACAGT  
CGGGGAAACACTTGGTCAAGAAGACACAAAACACTCTCAGAGACAGATGAGAAATTTCTGG  
GCCTTTAGGTAAATATTAGCTAAGAAAACCAAGGGGGAAATTGGAGGCAATTTTAAAAA  
AATAACGTGGACGCTATTAATGATTATCTTTGACGCTTGAAGTCATATAGCTCCTTGTAG  
TTTCTGTTAAGATCTCAAAGGAGGGTAACAGCAAGAAGCTCTGATTTTCACTGATTCTC

6632 TTCTCTCTCTCTTTCTCTTTCTCCCTCTCTCATCTTACCCTTTCCCCCACCAGGTCCC  
CGGCCAGTTGTGTATATGCAGCATGCCCTGTTTGCAGACAATGCCTACTGGCTTGAGAAT  
TATGCCAATGGAAGCCTTGGATTCCTTCTAGCAGATGCAGGTTATGATGTATGGATGGGA  
AACAGTCGGGGAAACACTTGGTCAAGAAGACACAAAACACTCTCAGAGACAGATGAGAAA  
TTCTGGGCCTTTAGGTAAATATTAGCTAAGAAAACCAAGGGGGAAATTGGAGGCAATTT  
[T,A]

FIG. 3R

AAAAAATAACGTGGACGCTATTAATGATTATCTTTGACGCTTGAAGTCATATAGCTCCT  
TGTAGTTTCTGTAAAGATCTCAAAGGAGGGTAACAGCAAGAAGCTCTGATTTTTCACTGA  
TTCTCCCACAAGCAAAGTATGGCATTTCACCAAGATCATTTTACATCCAATTCTGTGAA  
TTCTATGCATTAAGTATGTCCAAAGAGACAGCTCAGGAAATTATCATGACCAATGTGC  
ACATTCATTAGCCAATGTTTACTGAGTGGCTACTGTATGCGCTGTTCTAGGCCCGAAC

6763 AAGCCTTGGATTCTTCTAGCAGATGCAGGTTATGATGTATGGATGGGAAACAGTCGGGG  
AAACACTTGGTCAAGAAGACACAAAACACTCTCAGAGACAGATGAGAAATTCTGGGCCTT  
TAGGTAAATATTAGCTAAGAAAACCAAGGGGAAATTGGAGGCAATTTTAAAAAATAA  
CGTGGACGCTATTAATGATTATCTTTGACGCTTGAAGTCATATAGCTCCTTGTAGTTTCT  
GTAAAGATCTCAAAGGAGGGTAACAGCAAGAAGCTCTGATTTTTCACTGATTCTCCACA  
[A,G]  
GCAAAGTATGGCATTTCACCAAGATCATTTTACATCCAATTCTGTGAATTCTATGCATT  
AAAAGTATGTCCAAAGAGACAGCTCAGGAAATTATCATGACCAATGTGCACATTATTCA  
GCCAATGTTTACTGAGTGGCTACTGTATGCGCTGTTCTAGGCCCGAACATTCAAACAGG  
GAACAGACAACTCTGACCTCACAAAGCTTATGTTTATTAGTGATAATTTTACAAGTC  
ATTGCTCCTGGATTGCCAATCAACTGTGTAAAGATGATTTGGACCAGGACCTTATTGATT

6955 TAATGATTATCTTTGACGCTTGAAGTCATATAGCTCCTTGTAGTTTCTGTAAAGATCTCA  
AAGGAGGGTAACAGCAAGAAGCTCTGATTTTTCACTGATTCTCCACAAGCAAAGTATGG  
CATTTCAACAAGATCATTTTACATCCAATTCTGTGAATTCTATGCATTAAGTATGTC  
CAAAGAGACAGCTCAGGAAATTATCATGACCAATGTGCACATTATTAGCCAATGTTTA  
CTGAGTGGCTACTGTATGCGCTGTTCTAGGCCCGAACATTCAAACAGGGAACAGACAAA  
[- , T, C]  
TCTGACCTCACAAAGCTTATGTTTATTAGTGATAATTTTACAAGTCATTGCTCCTGGA  
TTGCCAATCAACTGTGTAAAGATGATTTGGACCAGGACCTTATTGATTTAGAGAACTGT  
GATTGATTTAGAGAACTGAGATCGCACATAGTACCATTTTACAGGAACTCCAATATTA  
GATTTTTAAACCTTGTTAATGGGCAATGAAGAAGAATCTTTTTGATATCTTGTTTCTT  
TTAATGGAAGAGTTTTCTGCTGTCACCAGAGGACAGGCTGATGCCTGCGATAGACTTTTC

7017 GGAGGGTAACAGCAAGAAGCTCTGATTTTTCACTGATTCTCCACAAGCAAAGTATGGCA  
TTTCAACAAGATCATTTTACATCCAATTCTGTGAATTCTATGCATTAAGTATGTCCA  
AAGAGACAGCTCAGGAAATTATCATGACCAATGTGCACATTATTAGCCAATGTTTACT  
GAGTGGCTACTGTATGCGCTGTTCTAGGCCCGAACATTCAAACAGGGAACAGACAACT  
CTGACCTCACAAAGCTTATGTTTATTAGTGATAATTTTACAAGTCATTGCTCCTGGAT  
[T,G]  
GCCAATCAACTGTGTAAAGATGATTTGGACCAGGACCTTATTGATTTAGAGAACTGTGA  
TTGATTTAGAGAACTGAGATCGCACATAGTACCATTTTACAGGAACTCCAATATTAGA  
TTTTTAAACCTTGTTAATGGGCAATGAAGAAGAATCTTTTTGATATCTTGTTTCTTT  
AATGGAAGAGTTTTCTGCTGTCACCAGAGGACAGGCTGATGCCTGCGATAGACTTTTCTT  
TCTCAGGCCTAAGCTCCCTGTTGGTTTGTAACCTGATGCTAGAACAGACTGTGATTTC

7151 GAAATTATCATGACCAATGTGCACATTATTAGCCAATGTTTACTGAGTGGCTACTGTA  
TGCGCTGTTCTAGGCCCGAACATTCAAACAGGGAACAGACAACTCTGACCTCACAAAG  
CTTATGTTTATTAGTGATAATTTTACAAGTCATTGCTCCTGGATTGCCAATCAACTGT  
GTAAAGATGATTTGGACCAGGACCTTATTGATTTAGAGAACTGTGATTGATTTAGAGAA

FIG. 3S



ACTGAGATCGCACATAGTACCATTTTCAGGAAAACCTCCAATATTAGATTTTTAAACCTT  
[G,T]

TTAATGGGCAATGAAGAAGAATCTTTTTTGATATCTTGTTTCTTTAATGGAAGAGTTTT  
CTGCTGTCACCAGAGGACAGGCTGATGCCTGCGATAGACTTTTCTTCTTCAGGCCTAAG  
CTCCCTGTTGGTTTGTAAACCTGATGCTAGAACAGACTGTGTATTCTATTACATTAATA  
AAACATTAGTACCCACTGAAAGTTTGAGAATAGTGGAGGAATAGAATAGAATGTTATAG  
TCTGAGTTCTTGGGCAGGGGCAAGCATCAGGAAATATTGAATCATTAGTCTTAGGAGGT

7308

CTCCTGGATTGCCAATCAACTGTGTAAAGATGATTTGGACCAGGACCTTATTGATTTAGA  
GAAACTGTGATTGATTTAGAGAACTGAGATCGCACATAGTACCATTTTCAGGAAAACCTC  
CAATATTAGATTTTTAAACCTTGTTAATGGGCAATGAAGAAGAATCTTTTTTGATATCT  
TGTTTCTTTAATGGAAGAGTTTTCTGCTGTCACCAGAGGACAGGCTGATGCCTGCGATA  
GACTTTTCTTCTTCAGGCCTAAGCTCCCTGTTGGTTTGTAAACCTGATGCTAGAACAGA  
[C,G]

TGTGTATTCCTATTACATTAATAAAACATTAGTACCCACTGAAAGTTTGAGAATAGTGG  
AGGAATAGAATAGAATGTTATAGTCTGAGTTCTTGGGCAGGGGCAAGCATCAGGAAATAT  
TGAATCATTAGTCTTAGGAGGTGTCACAACAATTCTCCTATTCTTGTAAGTCCCAATCT  
ATAGATTTCTCACATGTTCTTTAATAAACAGGCTTCTAGCTTATGGAATACCTGATTT  
GACTAAATGTTATATAGGCCCTTTTGTTCCCTCTGTCTGAAGAACAAAATACTAGTACTA

7321

AATCAACTGTGTAAAGATGATTTGGACCAGGACCTTATTGATTTAGAGAACTGTGATTG  
ATTTAGAGAACTGAGATCGCACATAGTACCATTTTCAGGAAAACCTCCAATATTAGATTT  
TTAAACCTTGTTAATGGGCAATGAAGAAGAATCTTTTTTGATATCTTGTTTCTTTAAT  
GGAAGAGTTTTCTGCTGTCACCAGAGGACAGGCTGATGCCTGCGATAGACTTTTCTTCT  
TCAGGCCTAAGCTCCCTGTTGGTTTGTAAACCTGATGCTAGAACAGACTGTGTATTCCTA  
[T,C]

TACATTAATAAAACATTAGTACCCACTGAAAGTTTGAGAATAGTGGAGGAATAGAATAG  
AATGTTATAGTCTGAGTTCTTGGGCAGGGGCAAGCATCAGGAAATATTGAATCATTAGTC  
TTTAGGAGGTGTCACAACAATTCTCCTATTCTTGTAAGTCCCAATCTATAGATTTCTCA  
CATGTTCTTTAATAAACAGGCTTCTAGCTTATGGAATACCTGATTTGACTAAATGTTAT  
ATAGGCCCTTTTGTTCCCTCTGTCTGAAGAACAAAATACTAGTACTATGGAATATTGGTA

7542

GCGATAGACTTTTCTTCTTCAGGCCTAAGCTCCCTGTTGGTTTGTAAACCTGATGCTAG  
AACAGACTGTGTATTCCTATTACATTAATAAAACATTAGTACCCACTGAAAGTTTGAGA  
ATAGTGGAGGAATAGAATAGAATGTTATAGTCTGAGTTCTTGGGCAGGGGCAAGCATCAG  
GAAATATTGAATCATTAGTCTTAGGAGGTGTCACAACAATTCTCCTATTCTTGTAAGTC  
CCAATCTATAGATTTCTCACATGTTCTTTAATAAACAGGCTTCTAGCTTATGGAATAC  
[C,T]

TGATTTGACTAAATGTTATATAGGCCCTTTTGTTCCCTCTGTCTGAAGAACAAAATACTA  
GTACTATGGAATATTGGTATATATTAAATATATATCTATATATCCATGTGGACAGGAATA  
CTACTACTAACAACATCTTACTGAGCACCCACTGGCAGCCAGAGTCGTTTCTTTCATACT  
ATTAACCCCGTTAGCAGCCCCGTAAACCAGGTACTACCCTGTTTATTTCCCAATGAGA  
AAACATAGGCTCAGAGCATTTAGTAATTTCTCAAGAGTTGCAAAGGCCATAAATAGTAG

8597

ATAAACTGGTCAGGAGAAATTGATTTTCATTGGACATTCCTTGGCACTACAATAGGTA  
TGTTTATGAGGGTCACTGTTAGGTGTGTTTTGAGGGTCAGTTTTCTCAGAGTCTTACAG

FIG. 3T

GAGTTCACCTTTATGTTGGAATAAAACAACCTGTTACTTATAGTGCCCTCAATTCCTGTC  
CTCTGCTGGGAATAACCCTAGTACTCTAAGTAGCTGTGAGCCTGCAGTGACAGACTATA  
TGTAGGGCAAACCTTTCCTGGGTCTCTGGTCACAGCAGCATATTGACTACGGTGATGCAA  
[T,C]  
TTCCCAGGAATAACATGTGTTCCAAATTCAAAGAAATAATTCCACAGAGTAAGTTTCTAG  
ATTCCCTCTGAGCTGAAAAAGTAAATTCATGCCATGGAATATGGCTGAAACATAATAA  
ATGTGCATCAATCATCTCTTTCTCACAACCCAAATGGGATTTTTAAAAAATAAAAGGGAA  
GGGCTTATACCTATATTTAAACAAATTGAAAAGGCATGGTTATATTTGTTGTGAGTTGG  
AACACACAAGCTTACTATAATAAATCAATTGAGCTTATCTATTCAAGTGTGTGATTTAGTA

8803 TAAGTAGCTGTGAGCCTGCAGTGACAGACTATATGTAGGGCAAACCTTTCCTGGGTCTC  
TGGTCACAGCAGCATATTGACTACGGTGATGCAATTTCCCAGGAATAACATGTGTTCCAA  
ATTCAAAGAAATAATTCCACAGAGTAAGTTTCTAGATTCCCTCTGAGCTGAAAAAGTAAA  
ATTCAATGCCATGGAATATGGCTGAAACATAATAAATGTGCATCAATCATCTCTTTCTCA  
CAACCCAAATGGGATTTTTAAAAAATAAAAGGGAAGGGCTTATACCTATATTTAAACAAA  
[C,T]  
TGAAAAGGCATGGTTATATTTGTTGTGAGTTGGAACACACAAGCTTACTATAATAAATC  
AATTGAGCTTATCTATTCAGTGTGTGATTTAGTATTTATGAAATAGCAAGTAAATGTAAG  
CACTATGTAGAAATTTCTAAAGTTTTTAAAGCTGACAACCTTACTTCTTAATTTACTTACT  
TTACTTAATTTACTTTACAATTTACTTTCCAGGTATTTTGAAAAGAAATCAATAATCTAG  
TTCCAAGTAAAAGTTGAAAAGGAACCCACACTAATAAAAGCTTTGAATTTGTCATTGAAC

9016 AAATGTGCATCAATCATCTCTTTCTCACAACCCAAATGGGATTTTTAAAAAATAAAAGGG  
AAGGGCTTATACCTATATTTAAACAAATTGAAAAGGCATGGTTATATTTGTTGTGAGTT  
GGAACACACAAGCTTACTATAATAAATCAATTGAGCTTATCTATTCAGTGTGTGATTTAG  
TATTTATGAAATAGCAAGTAAATGTAAGCACTATGTAGAAATTTCTAAAGTTTTTAAAGC  
TGACAACCTTACTTCTTAATTTACTTACTTTACTTAATTTACTTTACAATTTACTTTCCAG  
[G,A]  
TATTTTGAAAAGAAATCAATAATCTAGTTCCAAGTAAAAGTTGAAAAGGAACCCACACTAA  
TAAAAGCTTTGAATTTGTCATTGAACCTCCACTAAAGTTTCCAATTTTAAAGAGAATAAAT  
CATGTGAAAGTGCAATATTTTCAGTTTAGGGAAATATTTTCATTATCACCCTATCATCAG  
TAACAAACATATATTCATTAGTATTTTAGATTGACAGGCACTTTCCAAGCTCAGAACAGG  
CAGTTAGCATCAGTCAGCATATACTAAAAAAGTATCAAAGAACTCATAGGAGATCAAAAA

9967 GTTTCATTTAGGACATAAATATTTTTAGTGACTGTTGTTTGCATTTTGGACAGAGCAATT  
TCTGTTATGTAAGGAGCACCCACTCTTTGTAGGACATTTAGTAGGTCCCAGCCATTAAA  
CAGGGCTCTGCAGTCAGCGTGACCCTCAAAAATCTCACCTCCACACATTTCCAAACACCC  
TCTGGGGAAGTACTATTCTGATTCAGAGTCTTTTTATCAATTGTTCAAGTCAATTATTTT  
AGTTCTTCTTTTTCTGGCCAAGACAGTTTAAATGTTCCAACAAGTGTTCAGTACACACA  
[T,C]  
ACACACACACACACACACACACACACACACACATGCTAGTGGAGGCCAGGAAGGG  
ACCTCTGGAAACCAAATTATATGGATATTTCTCCCTAGCCTACCCAGTGTGTGCTAATCT  
CCATCCTCAGATATACAAAGGGGTGCAATGCTACTGCTGAAAGAGCAAAGCAAATGGA  
GATGCCTGGTCCTTACTGGGCCATCGTGGATGCTAGGGAAAGCCCCTTTCTTTTTGGAAA  
CAGGGAAGAGTCTAGAGGGTTGAAAAACACCCAGTAAGACACTGGGAGCAGTGAAATTTT

FIG. 3U

- 10008 CATT TTTGGACAGAGCAATTTCTGTTATGTAAGGAGCACCCACTCTTTGTAGGACATTTAG  
TAGGTCCCAGCCCATTA AACAGGGCTCTGCAGTCAGCGTGACCCTCAAAAATCTCACCTC  
CACACATTTCCAAACACCCTCTGGGGAAGTACTATTCTTGATTGAGAGTCTTTTATCAA  
TTGTTCAAGTCAATTATTTCAAGTTCTTCTTTTCTGGCCAAGACAGTTTAAATGTTCCAAC  
AAGTGTTCAGTACACACATACACACACACACACACACACACACACACACACATGC  
[C,T]  
AGTGGAGGCCAGGAAGGGACCTCTGGAACCAAATTATATGGATATTCTCCCTAGCCTA  
CCCAGTGTGTGCTAATCTCCATCCTCACAGATATACAAAGGGGTGCAATGCTACTGCTG  
AAAGAGCAAAGCAAATGGAGATGCCTGGTCCTTACTGGGCCATCGTGGATGCTAGGGAAA  
GCCCTTTCTTTTGGAAACAGGGAAGAGTCTAGAGGGTTGAAAAACCCAGTAAGACA  
CTGGGAGCAGTGAAATTTCAATCCATAGTGAGAAAGAAAACCTGTTAGAATAACTGGGTG
- 10363 AGCCTACCCAGTGTGTGCTAATCTCCATCCTCACAGATATACAAAGGGGTGCAATGCTA  
CTGCTGAAAGAGCAAAGCAAATGGAGATGCCTGGTCCTTACTGGGCCATCGTGGATGCTA  
GGGAAAGCCCCTTTCTTTTGGAAACAGGGAAGAGTCTAGAGGGTTGAAAAACCCAGT  
AAGACACTGGGAGCAGTGAAATTTCAATCCATAGTGAGAAAGAAAACCTGTTAGAATAAC  
TGGGTGATGCTGCAGAAAGAAATCAATTCACCTCCTGTGACTGATTATTTGCTTCTGGAA  
[G,A]  
CTCTGTGATTCAATCTGGCATCTCAGAGTTAGGGATGAAATGAGAATGTTGCCAGCATTT  
ACCCCATGCTTGGGAAGTTTACACAGCAGTAGCTACTCCAGCAGCTTAACCATCACCTTT  
CCCCTGCCAACTACTCCATTTCCCCCAATCAAGTCAAATGTCCATAAATAGAATAAAAT  
AAAATTGGAGACTTGAGAGCAGAGAAGACTGAAGGCAGATTATCTTTATAGAATAACTCA  
GAAGACTTCCAATTCATCCCCAGTATGATCACGATAGAAGGAAAAAATGACTAAGCAGAG
- 10684 TCTCAGAGTTAGGGATGAAATGAGAATGTTGCCAGCATTTACCCCATGCTTGGGAAGTTT  
ACACAGCAGTAGCTACTCCAGCAGCTTAACCATCACCTTTCCCCTGCCAACTACTCCATT  
TCCCCCAATCAAGTCAAATGTCCATAAATAGAATAAAATAAAATTGGAGACTTGAGAGC  
AGAGAAGACTGAAGGCAGATTATCTTTATAGAATAACTCAGAAGACTTCCAATTCATCCC  
CAGTATGATCACGATAGAAGGAAAAAATGACTAAGCAGAGCCCCAATTTTGTAGAAACA  
[T,C]  
TGCGTAAGTATTTATTTTACAAGATTGTCTTATCTCCTGTTCTCTCAGGGTTTGTAGCC  
TTTTCCACCATGCCTGAACTGGCACAAGAATCAAATGAATTTTGCCTTGGGTCTACG  
ATCTCATTCAAATATCCACGGGCATTTTACCAGGTTTTTCTACTTCAAATTCATA  
ATCAAGGTAGGCTCCTTTCAACAAATGTACCTGAGGATCTATTTTGGATCATAAATCC  
TTATTATTTTCAAATCTACTGTAAAGTAAAGTAGGAAATTTAGATAAAATCTATAGAAC
- 11177 TCCTTTCAACAAAATGTACCTGAGGATCTATTTTGGATCATAAATCCTTATTATTTTCA  
AATCTACTGTAAAGTAAAGTAGGAAATTTAGATAAAATCTATAGAACTTAGACTCTGTG  
GGTATGTGCTTGTGTATGTGTGCCCTGCGTGTGCGCATGTCTGTGCCATAGTATCTGCA  
GGTTCTGTAATACAATTTACTATACAAGGTATCAGCAGGCTGAGTATATGTCAGAAATTT  
CTAGCTGAACTGAGTGCTATATGACAACAAGGATTTTCTTGTTTTCCCAAGTGTTTTTT  
[G,T]  
TTCCATTTAGTCAGGTAGGTCAATGAATTCACATTGCCCAAATGAAAGACACTTCAAGTT  
ACCCATAATCACTGATGTGTCCAATTTTGCATTAGAAAAACCTGATTAATATATTCCTT  
CCAATATGGAACTTGCCCTAATAACTAAAGCTAAGATTCCAAAGCCTAAATGTATTACA  
GCTCAAGTATTAATTCAAATATTTATTGGTTATTTTTCAGGAGTTGAAAAAGTCATTTGG

FIG. 3V

TTGCCAATTGTGGATTTGGGATTTTATCTATTAAGGGTTTTTTTTTTTTTCTCTTGC

12345 TTTAAGTCCCATATCCTGCTCTTTTCTTCCGTCAGTTTCCCCAGAAGCTCCAAGACCCC  
ACCAGGAATCCCCATCCAAGTTTACTTTCCCAACTCCTGGAAGTTTCAATTGTGCTGCCT  
TTGTGACATTATCATATCTTTTCTGTTCAATGGTTGCTTCTCTTTGGCTCACTGTTCTCT  
ACTTTTCAGCCTGAGAGCTGGCTAATCTGGGACAGTACTCGAATGCAGTGTACACATGGG  
TAACATGGAAAACCCCGATTTTCCCTTATATTCAAGGTATTATTTGACCTTAAGAAAAAC  
[T,C]  
GTTTTACATTTCATACCAATTAATGAGAAAAAATATTGGCAAGCACTGACTGGGCAGAA  
TACAGGAAGCTTCACTATGGAGAAGTGAATTTGGGATTGAGGGCCTTTATTGCAATCTC  
CTTGTAATAATATTTGATACTCTTCCTCATCTGGAGACACATTCTAAGTAACTTTTCC  
TGAATAATTTGGTCTCCTTGACTGAATCAGTAAGTACAAATAGATCCCCAAGCATGGCTC  
TTTCTAGAATGAAAGAAATGTCAAGAAGTCTGAAGATGATTCTTGAATTTTGGTTTTTT

12349 AGTCCCATATCCTGCTCTTTTCTTCCGTCAGTTTCCCCAGAAGCTCCAAGACCCCACCA  
GGAATCCCCATCCAAGTTTACTTTCCCAACTCCTGGAAGTTTCAATTGTGCTGCCTTTGT  
GACATTATCATATCTTTTCTGTTCAATGGTTGCTTCTCTTTGGCTCACTGTTCTCTACTT  
TTCAGCCTGAGAGCTGGCTAATCTGGGACAGTACTCGAATGCAGTGTACACATGGGTAAC  
ATGGAAAACCCCGATTTTCCCTTATATTCAAGGTATTATTTGACCTTAAGAAAAACTGTT  
[C,T]  
TACATTTCATACCAATTAATGAGAAAAAATATTGGCAAGCACTGACTGGGCAGAATACA  
GGGAAGCTTCACTATGGAGAAGTGAATTTGGGATTGAGGGCCTTTATTGCAATCTCCTTG  
TAAATAATATTTGATACTCTTCCTCATCTGGAGACACATTCTAAGTAACTTTTCTGAA  
TAATTTGGTCTCCTTGACTGAATCAGTAAGTACAAATAGATCCCCAAGCATGGCTCTTTC  
CTAGAATGAAAGAAATGTCAAGAAGTCTGAAGATGATTCTTGAATTTTGGTTTTTTGCTA

13115 TAGAAGATAAGAAAACGAAGATAGCTTCTACCAAAATCTGCAACAATAAGATACTCTGGT  
TGATATGTAGCGAATTTATGTCCTTATGGGCTGGATCCAACAAGAAAAATATGAATCAGG  
TATGTATGATAATTATAGGGCCATTTGATACCTTAAGAAATTCAGCTTTCTTTTGACTC  
ATTTTGATATATCTATTTACTGTATAAATTCATATGGTATTCCAAACCCTTAAGACAGA  
TTTTTTTTTGCTTTTAAAAATGTTTATGGGTATATAATAGTTGTACATATTTATGAGACA  
[C,T]  
ATATATTTTGATATAAGCATACAATGTGTAATGACCAAATCAGGGTAATTGGGATATCCA  
TCACCTCAAGCATTTATCATTCTTTTGTAGAGACATTCTAATTTGACTCTTCTAGTT  
ATTTTGAAATATACAATGAATTATTGTTAACTATAGTCATCCTATTGTGCATGCCAGACT  
TTAGTCCTTCTAACGGTATTTTGGTACCCATTAACCAATGCCTCTTTATCCTTCCCCAC  
CCCTACTACCTTTCCAGCCTCTGGTAACCATCATTCTTCTCACTATCTCTATAAGGTCA

13354 ATTTTTTTTTGCTTTTAAAAATGTTTATGGGTATATAATAGTTGTACATATTTATGAGAC  
ACATATATTTTGATATAAGCATACAATGTGTAATGACCAAATCAGGGTAATTGGGATATC  
CATCACCTCAAGCATTTATCATTCTTTTGTAGAGACATTCTAATTTGACTCTTCTAG  
TTATTTTGAAATATACAATGAATTATTGTTAACTATAGTCATCCTATTGTGCATGCCAGA  
CTTTAGTCCTTCTAACGGTATTTTGGTACCCATTAACCAATGCCTCTTTATCCTTCCCC  
[T,A]  
CCCCTACTACCTTTCCAGCCTCTGGTAACCATCATTCTTCTCACTATCTCTATAAGGTC  
AGTTTTTTTTTAACTCCCCTATATGAGTGAGAACATGCAGTATTTGTCTTTTGTGCCT

FIG. 3W

GGCTTATTTCACTTAATGTAATGTTCTCTAATTTATCCACATTATTGCAAATGACATGA  
TTTCATTCTTCTTATGGCTGTCTATATGTACCACATTTTATTTATCCACTCATCTGTTGA  
TGGACACTTAGGCTGATTTTCATATCTTGGTCATTGTGAATAGTGCTGTACTAAACATGGG

13373 AATGTTTATGGGTATATAATAGTTGTACATATTTATGAGACACATATATTTTGATATAAG  
CATACAATGTGTAATGACCAAATCAGGGTAATTGGGATATCCATCACCTCAAGCATTAT  
CATTTCTTTTTGTTAGAGACATTCTAATTTGACTCTTCTAGTTATTTTGAAATATACAAT  
GAATTATTGTTAACTATAGTCATCCTATTGTGCATGCCAGACTTTAGTCCTTCTAACGGT  
ATTTTGGTACCCATTAACCAATGCCTCTTTATCCTTCCCCACCCCTACTACCTTTCCCA  
[C,G]  
CCTCTGGTAACCATCATTCTTCTCACTATCTCTATAAGGTCAGTTTTTTTTTAAACTCCC  
CTATATGAGTGAGAACATGCAGTATTTGTCTTTTTGTGCCTGGCTTATTTCACTTAATGT  
AATGTTCTCTAATTTATCCACATTATTGCAAATGACATGATTTTATTCTTCTTATGGCT  
GTCTATATGTACCACATTTTATTTATCCACTCATCTGTTGATGGACACTTAGGCTGATTT  
CATATCTTGGTCATTGTGAATAGTGCTGTACTAAACATGGGGTGCAGATGTCTCTTCCA

14677 AGAGATAGAGATCTAATTTTATTCTTCTGCATATGGATATCTAGTTTTCCAGCATCATT  
TCTTGTGGAAATTGTCCTTTGCCCAATGTATGTTCTTGATGCCTTTGTTGAAAATTAGTT  
GACTATAAATGTGTGGATTTATTTGTGGGTTCTTTATTCTGTTCCATTGGTCTATGTGTC  
TGTTTTTATGCCAGTATCATGCAGTTTTGATTATTACAGGTTTGTAGTATAATTTGAAGT  
CAGGTCATGTGATGCCTCCAGCTTTGTTCTTTTTCTCAGAATCTTATATTTAGAAAAAC  
[C,G]  
TAAAGACTCCAACAAAAAACCTGCTAGAACTGATAAACAAATTCATTAAATTTGCAGGAT  
ACAACATCAACATACAAAATTCAGCAGCATTTCAATATGCCAAGAGCAAATAATCTTAAA  
AAAAAGAAAGAAAAAACAAGAAATAATCCCATTTATAATAGCTACAAATAAAATAAA  
ACACCTAGGAATAAACCATACCAAAGAAGTGAAAGATTTCTACAATGAAAACATATAAAAC  
ACTGATGAAAGAAATTGAAAATGACATTAAAAAATGGAAAGGTATTCCATGTTTCATGGAT

14734 ATTTCTTGTGGAAATTGTCCTTTGCCCAATGTATGTTCTTGATGCCTTTGTTGAAAATTA  
GTTGACTATAAATGTGTGGATTTATTTGTGGGTTCTTTATTCTGTTCCATTGGTCTATGT  
GTCTGTTTTTATGCCAGTATCATGCAGTTTTGATTATTACAGGTTTGTAGTATAATTTGA  
AGTCAGGTCATGTGATGCCTCCAGCTTTGTTCTTTTTCTCAGAATCTTATATTTAGAAA  
AACGTAAAGACTCCAACAAAAAACCTGCTAGAACTGATAAACAAATTCATTAAATTTGCA  
[G,A]  
GATACAACATCAACATACAAAATTCAGCAGCATTTCAATATGCCAAGAGCAAATAATCTT  
AAAAAAGAAAGAAAAAACAAGAAATAATCCCATTTATAATAGCTACAAATAAAAT  
AAAACACCTAGGAATAAACCATACCAAAGAAGTGAAAGATTTCTACAATGAAAACATATAA  
AACACTGATGAAAGAAATTGAAAATGACATTAAAAAATGGAAAGGTATTCCATGTTTCATG  
GATTGCAAGAATCAATATTGTTAAAATGTCCATATGATCCAAAAAATCTACAGATTCAA

14747 ATTGTCCTTTGCCCAATGTATGTTCTTGATGCCTTTGTTGAAAATTAGTTGACTATAAAT  
GTGTGGATTTATTTGTGGGTTCTTTATTCTGTTCCATTGGTCTATGTGTCTGTTTTATG  
CCAGTATCATGCAGTTTTGATTATTACAGGTTTGTAGTATAATTTGAAGTCAGGTCATGT  
GATGCCTCCAGCTTTGTTCTTTTTCTCAGAATCTTATATTTAGAAAAACGTAAAGACTC  
CAACAAAAAACCTGCTAGAACTGATAAACAAATTCATTAAATTTGCAGGATACAACATCA  
[A,G]

FIG. 3X

CATACAAAATTGAGCAGCATTTCATATGCCAAGAGCAAATAATCTTAAAAAAGAAAG  
AAAAAACAAGAAATAATCCCATTTATAATAGCTACAAATAAAACACCTAGGA  
ATAAACCATACCAAAGAAGTGAAAGATTTCTACAATGAAACTATAAAACACTGATGAA  
GAAATTGAAATGACATTAATAATGGAAGGTATTCCATGTTTCATGGATTGCAAGAATC  
AATATTGTTAAATGTCCATATGATCCAAACAATCTACAGATTCAATGCAATCCCTATC

14808 TGTGGATTTATTTGTGGGTTCTTTATTCTGTTCCATTGGTCTATGTGTCTGTTTTATGC  
CAGTATCATGCAGTTTTGATTATTACAGTTTTGTAGTATAATTTGAAGTCAGGTCATGTG  
ATGCCTCCAGCTTTGTTCTTTTTCTCAGAATCTTATTTAGAAAAACGTAAAGACTCC  
AACAAAAACCTGCTAGAACTGATAAACAATTCATTAAATTTGCAGGATACAACATCAA  
CATACAAAATTGAGCAGCATTTCATATGCCAAGAGCAAATAATCTTAAAAAAGAAAG  
[-,A]  
AAAAAACAAGAAATAATCCCATTTATAATAGCTACAAATAAAACACCTAGGAA  
TAAACCATACCAAAGAAGTGAAAGATTTCTACAATGAAACTATAAAACACTGATGAAAG  
AAATTGAAATGACATTAATAATGGAAGGTATTCCATGTTTCATGGATTGCAAGAATCA  
ATATTGTTAAATGTCCATATGATCCAAACAATCTACAGATTCAATGCAATCCCTATCA  
AAATACCAATGACATTCTTCATTGAAATAAAAAAAGCCTAAAATTTAAGTGAACCAT

15086 AATAATCTTAAAAAAGAAAGAAAAAACAAGAAATAATCCCATTTATAATAGCTAC  
AAATAAAATAAAACACCTAGGAATAAACCATACCAAAGAAGTGAAAGATTTCTACAATGA  
AACTATAAAACACTGATGAAAGAAATTGAAATGACATTAATAATGGAAGGTATTCC  
ATGTTTCATGGATTGCAAGAATCAATATTGTTAAATGTCCATATGATCCAAACAATCTA  
CAGATTCAATGCAATCCCTATCAAATACCAATGACATTCTTCATTGAAATAAAAAA  
[-,A,G]  
CCTAAAATTTAAGTGAACCATGAAGGTAGATGTCTGCTATACATAGAAGATTAAGTACT  
CAACAAACCTTGAATATGAAGACTGGGAAGTGAATAGGCAGCTTCACTCTTCTATTCCC  
TGGTGAATTTAGGAGAATGGATGTTTTATAATGGGTAGCAGTTTCTTACATGTTCTCAA  
TCAGCCATAACTTACTACAGTCAATTTGAATTTATTGCATTTGAATATATTGGATTAAAA  
ATAAATCCTAAAAAAGGAGAGAAGCACATATAAACCTGCGTCTTATTTTCATGTGTTCT

15414 TAGATGTCTGCTATACATAGAAGATTAAGTACTCAACAAACCTTGAATATGAAGACTGGG  
GAAGTGAATAGGCAGCTTCACTCTTCTATTCCCTGGTGAAATTTAGGAGAATGGATGTTT  
TATAATGGGTAGCAGTTTCTTACATGTTCTCAATCAGCCATAACTTACTACAGTCAATTT  
GAATTTATTGCATTTGAATATATTGGATTAAAAATAAATCCTAAAAAAGGAGAGAAGCA  
CATATAAACCTGCGTCTTATTTTCATGTGTTCTTTCTTTGTGGGTGACTTTTGTGTTGAA  
[A,G]  
TAAACCTGCAAAATAACAGGACAGGGTGGAAGGGAGATGGGATCCCCTCTTTATGAAGA  
AGCAGCAGTCTGTTTTATCACCTCTTCATTTTCTGTTATTGAGAATTCAAGAAGAAGGA  
GGAGGAAGAGTTCACATCCACAGACTGGTGTGGTTGAATAGTTGTCTCTACTGTATTCCA  
AATAGCAGCCAATGAGGCTGTTACAGTGAAGCCAGTCCCAAGATAATTGTTCTGTACCCC  
TATTCTCTAAGAAGCTAAATTGTGTTAGACTGAAACCCATAAGGAACCATGTTCAAAGT

15722 TGCAAAATAACAGGACAGGGTGGAAGGGAGATGGGATCCCCTCTTTATGAAGAAGCAGCA  
GTCCTGTTTTATCACCTCTTCATTTTCTGTTATTGAGAATTCAAGAAGAAGGAGGAGGAA  
GAGTTCACATCCACAGACTGGTGTGGTTGAATAGTTGTCTCTACTGTATTCCAAATAGCA  
GCCAATGAGGCTGTTACAGTGAAGCCAGTCCCAAGATAATTGTTCTGTACCCCTATTCTC

FIG. 3Y

TAAGAAGCTAAATTGTGTTAGACTGAAACCCATAAGGAACCATTTGTTCAAAGTTGGCTTG  
[T,C]  
TCAAAAGTAAAGATTTTTAATAGTTTCTCTTAATTAGATTATTTTCTAAGACATAGAATT  
ATGATTACTATTTTATCTCTATAATTTTCATCTCTATAACGTTTACAAATACTGAAATAA  
CCTTTGGAAAAAATTGGCTTTTAGCTTTACTTTTGCAATATTTTATTTTATCCCCATAAA  
AGCCTAGGAAATTGGTACTATGACTTTTAGTATGTTCATTTAATAGATGAAAACACAGAA  
ACTCAAAGATGTTAAATATGGTGGCCAAGTTCACAAAGCTGATCATTAAACAACAGGG

15861 GGTGTGGTTGAATAGTTGTCTCTACTGTATTCCAAATAGCAGCCAATGAGGCTGTTACAG  
TGAAGCCAGTCCCAAGATAATTGTTCTGTACCCCTATTCTCTAAGAAGCTAAATTGTGTT  
AGACTGAAACCCATAAGGAACCATTTGTTCAAAGTTGGCTTGTTCAAAAGTAAAGATTTTT  
AATAGTTTCTCTTAATTAGATTATTTTCTAAGACATAGAATTATGATTACTATTTTATCT  
CTATAATTTTCATCTCTATAACGTTTACAAATACTGAAATAACCTTTGGAAAAAATTGGC  
[T,C]  
TTTAGCTTTACTTTTGCAATATTTTATTTTATCCCCATAAAAGCCTAGGAAATTGGTACT  
ATGACTTTTAGTATGTTCATTTAATAGATGAAAACACAGAAACTCAAAGATGTTAAATAT  
GGTGGCCAAGTTCACAAAGCTGATCATTAAACAACAACAGGGCCTGAACTCCTGGTTTTCT  
GATTTAATCTGTGACAGTGCACCTGGGTGCGCATGCATGCATCACCCCCACACTTGCACA  
TAGAACCTTTCCTAGTTGGCTTTGCTCCATGATGACCATTACTGTTCCCTTCTACTTCAA

16264 CTCAAAGATGTTAAATATGGTGGCCAAGTTCACAAAGCTGATCATTAAACAACAACAGGGC  
CTGAACTCCTGGTTTTCTGATTTAATCTGTGACAGTGCACCTGGGTGCGCATGCATGCAT  
CACCCCCACACTTGCACATAGAACCCTTTCCTAGTTGGCTTTGCTCCATGATGACCATTAC  
TGTTCCCTTCTACTTCAAATAAGCAAATTATCCTACAGATTCAGAGCTGGTACAGGTGTG  
CTGTCAAGCAGCCCATTCCATTAGTCAGCTTGTGGTTCACTCACATTAAAGTATTGACCT  
[A,T]  
AATGGTATATTTATCTAGATAATTCTACCTTGTTATTTTCAAAGCCCCAGTCTTGTTTGC  
TAATTCGTGTCATCATTTTTCTCTGATTCTGAAAGGCCAAAATTTGTTGGGCAATTGCTG  
TAATATGAGTTTTATCTCCTTTAGAGTGAATGGATGTGTATATGTCACATGCTCCCACT  
GGTTCATCAGTACACAACATTCTGCATATAAAACAGGTAGAGTCTTAGTCATGGAAAACC  
ATTCCAATCCTTATTTTCAATATATTTAAAAAGACAGAATTGACCCTGTTAACAGGCCTA

16314 ACAACAGGGCCTGAACTCCTGGTTTTCTGATTTAATCTGTGACAGTGCACCTGGGTGCGC  
ATGCATGCATCACCCCCACACTTGCACATAGAACCCTTTCCTAGTTGGCTTTGCTCCATGA  
TGACCATTACTGTTCCCTTCTACTTCAAATAAGCAAATTATCCTACAGATTCAGAGCTGG  
TACAGGTGTGCTGTCAAGCAGCCCATTCCATTAGTCAGCTTGTGGTTCACTCACATTAAA  
GTATTGACCTAAATGGTATATTTATCTAGATAATTCTACCTTGTTATTTTCAAAGCCCCA  
[G,A]  
TCTTGTTTGCTAATTCTGTGTCATCATTTTTCTCTGATTCTGAAAGGCCAAAATTTGTTGG  
GCAATTGCTGTAATATGAGTTTTATCTCCTTTAGAGTGAATGGATGTGTATATGTCACA  
TGCTCCCACTGGTTCATCAGTACACAACATTCTGCATATAAAACAGGTAGAGTCTTAGTC  
ATGGAAAACCATTCGAATCCTTATTTTCAATATATTTAAAAAGACAGAATTGACCCTGTT  
AACAGGCCTACCCTAAGAATCTTAAGAGCTTGCTCCAGTTTGTCTTGCTGCCTTCTGT

16877 TAAGAGCTTGCTTCCAGTTTGTCTTGCTGCCTTCTGTATGCCTTGATTTCCCTGGAATT  
TAAGAGAAAGGATGTTATGGTACAGACCAAGTAGATGACATAAATGAACACCACCTTAA

FIG. 3Z

TCAGAGTTTTAAAAATAGGCCCTGAACTGAAGCAAGAGGTAACTAGGGAAGCCTCAGGA  
GAACTGAGACTTCTCCAGAGAGAAGTATCTGGGATTTAACTTCTTTCTAATGAGGCTTGG  
TTTTCCATGAACTTTTCTTTAAACCAAGGGGGGTATTGCTCATCTTTCTGTTGAGCCCC  
[A,G]

TTTGTGATAATTGTAAAATGGGTGGTTACATCCTTCTGGTGATCTAGGAGCCCTATTTTC  
GTCCTAGCATACAGCATTTTCTAAAATTTGCTGTTAGCTTTCATGATTCTTACCCTAAC  
TATTCTTTTTCTAAAAACATTTGTTTCAGCTTACCCTCTGATGAATTCAGAGCTTAT  
GACTGGGGAAATGACGCTGATAATATGAAACATTACAATCAGGTGAGCTATTTACAGTAA  
CCCCAGCATGCTGATTTTGATAAATTATAATAAAAAATTATTTGAGGGTGGAAAGACTCC

16966 AGTAGATGACATAAATGAACACCACCTTAAATCAGAGTTTTAAAAATAGGCCCTGAACTG  
AAGCAAGAGGTAACTAGGGAAGCCTCAGGAGAAGTATCTCCAGAGAGAAGTATC  
TGGGATTTAACTTCTTTCTAATGAGGCTTGGTTTTCCATGAACTTTTCTTTAAACCAAG  
GGGGGTATTGCTCATCTTTCTGTTGAGCCCCATTTGTCATAATTGTAAAATGGGTGGTTA  
CATCCTTCTGGTGATCTAGGAGCCCTATTTTCGTCCTAGCATACAGCATTTTCTAAAAT  
[T,G]

TGCTGTTAGCTTTCATGATTCTTACCCTAACTATTCTTTTTCTAAAAACATTTGTTTCA  
GCTTTACCCTCTGATGAATTCAGAGCTTATGACTGGGGAAATGACGCTGATAATATGAA  
ACATTACAATCAGGTGAGCTATTTACAGTAACCCAGCATGCTGATTTTGATAAATTATA  
ATAAAAAATTATTTGAGGGTGGAAAGACTCCTACCTGTCATTTGGTGGCATTATATACTGA  
TAGAACTTTTTTTTAAAAAAATTTTAATTTTAATTTAATTTATTTTCAGAAAATTTATAA

17147 GGGGTATTGCTCATCTTTCTGTTGAGCCCCATTTGTCATAATTGTAAAATGGGTGGTTAC  
ATCCTTCTGGTGATCTAGGAGCCCTATTTTCGTCCTAGCATACAGCATTTTCTAAAAT  
TGCTGTTAGCTTTCATGATTCTTACCCTAACTATTCTTTTTCTAAAAACATTTGTTTCA  
GCTTTACCCTCTGATGAATTCAGAGCTTATGACTGGGGAAATGACGCTGATAATATGAA  
ACATTACAATCAGGTGAGCTATTTACAGTAACCCAGCATGCTGATTTTGATAAATTATA  
[A,G]

TAAAAAATTATTTGAGGGTGGAAAGACTCCTACCTGTCATTTGGTGGCATTATATACTGAT  
AGAACTTTTTTTTAAAAAAATTTTAATTTTAATTTAATTTATTTTCAGAAAATTTATAAA  
TTAAAGAAGCATATACAAAGAACTTACATCATGTGTAATCCTTCCATCCAGAGATAACT  
AGATGTACTAACATTTTGGTGTATTTATTCCAATTTTCTCAGTATTATATTGCTTTTAGA  
CAACTTTTAATCTTTCTATTTTACTTAAGCTATAGTAAGAGATAACTAATATAACTGAGG

17219 ATCTAGGAGCCCTATTTTCGTCCTAGCATACAGCATTTTCTAAAATTTGCTGTTAGCTT  
TCATGATTCTTACCCTAACTATTCTTTTTCTAAAAACATTTGTTTCAGCTTACCCTC  
TGATGAATTCAGAGCTTATGACTGGGGAAATGACGCTGATAATATGAAACATTACAATCA  
GGTGAGCTATTTACAGTAACCCAGCATGCTGATTTTGATAAATTATAATAAAAAATTAT  
TTGAGGGTGGAAAGACTCCTACCTGTCATTTGGTGGCATTATATACTGATAGAACTTTTTT  
[T,C]

TAAAAAAATTTTAATTTTAATTTTAATTTATTTTCAGAAAATTTATAAATTAAAGAAGCAT  
ATACAAAGAACTTACATCATGTGTAATCCTTCCATCCAGAGATAACTAGATGTACTAAC  
ATTTTGGTGTATTTATTCCAATTTTCTCAGTATTATATTGCTTTTAGACAACCTTTAATC  
TTTCTATTTTACTTAAGCTATAGTAAGAGATAACTAATATAACTGAGGGATTTTAAATG  
CATTTTAAATGGCTACATAATAGAAATTATTTCATAAAAATCTTTACAGCATAAATGAAT

FIG. 3AA



18628 AAAATGAAACAAAATCAACACGCACATTCAAGATCATTATGGTCAAGTAAAGTATGT  
GAGAGTGTTAATGTCCTTAGAATTTGGCCACAGTTAGCTGGTCCTACTCTGCTCCAAGCC  
GGTCCTATTTTGTGAATTAATCTCATTGATGCCAATTTTATTACATTCTCTCAAAAA  
ACTAGTCTCAACAGTTTGCTCTCTCCTCAAGTTCACAGCATTATCTCTGCTATATCTATA  
TTTTATTGAGTATAAGAGAATTAACCCATGTAAGCTCCATGAGGGTAGGGATTTCTCATC  
[A,G]  
TTTTGTTACCAGTGTTTTCTCATCTTGAAGAGTACATGACAATTACTGGGCTCCCAGTA  
TCTATGTGTTGCATTAATGAAATTTCTTAACTTTAATCTACCTCAAAATGTCTCTATCTT  
CTTGATTCTCTCCTTCCTTTCTCTATCAGAAAATGATGGTCCTCTTATTTTCCAAGTTAT  
TCCGGTCCTGTGCCCTTGATCCCATCTCTTCTCACTTCCCCTTCCTTCCTGCCTCCATTC  
TCCTGTCCCTTATGAAAAACAAGCAAGACCATCAATTCTATCAAGTTATCATTATGTCAC

18655 TCAAGATCATTATGGTCAAGTAAAGTATGTGAGAGTGTTAATGTCCTTAGAATTTGG  
CCACAGTTAGCTGGTCCTACTCTGCTCCAAGCCGGTCCTATTTTGTGAATTAATCTCATT  
TGATGCCAATTTTATTACATTCTCTCAAAAACTAGTCTCAACAGTTTGCTCTCTCCT  
CAAGTTCACAGCATTATCTCTGCTATATCTATATTTATTGAGTATAAGAGAATTAACCC  
ATGTAAGCTCCATGAGGGTAGGGATTTCTCATCGTTTTGTTACCAGTGTTTTCTCATCT  
[T,G]  
GAAGAGTACATGACAATTACTGGGCTCCCAGTATCTATGTGTTGCATTAATGAAATTTCT  
TAACTTTAATCTACCTCAAAATGTCTCTATCTTCTTGATTCTCTCCTTCCTTTCTCTATC  
AGAAAATGATGGTCCTCTTATTTTCCAAGTTATTCCGGTCCTGTGCCCTTGATCCCATCT  
CTTCTCACTTCCCCTTCCTTCCTGCCTCCATTCTCCTGTCCCTTATGAAAAACAAGCAAG  
ACCATCAATTCTATCAAGTTATCATTATGTCACCTCTGTTCTTATCAACATATTTTATGTA

18984 CAGTATCTATGTGTTGCATTAATGAAATTTCTTAACTTTAATCTACCTCAAAATGTCTCT  
ATCTTCTTGATTCTCTCCTTCCTTTCTCTATCAGAAAATGATGGTCCTCTTATTTTCAA  
GTTATTCCGGTCCTGTGCCCTTGATCCCATCTCTTCTCACTTCCCCTTCCTTCCTGCCTC  
CATTCTCCTGTCCCTTATGAAAAACAAGCAAGACCATCAATTCTATCAAGTTATCATTAT  
GTCACCTCTGTTCTTATCAACATATTTTATGATTGAAGAGGGCTTCTTCTACTTACTCCT  
[G,T]  
AACCTTGACAATGTAGTTTAGGTCTTCATCTTTTATCATAGCTACCTTATTTAAAGTC  
ACCCATGGCTTTTAATTGCCAAATTCAATGGCCTATCTTCACCTTTTGAAATGTGTTATG  
TTCGTTACCACAGTCTCCTTGAAACTCAGTCCCCTGACTTGGACTTCATAACACAATGA  
TTTCTGATTTTCTTCTGTTTGTGATTGTTCTTTTGTCCCAGGCACTGGCTACTCCACC  
TTCCACCTCTCTGAAATCATTAGCATTCCCCAAGGATTCTTCAAACTCTCTTTCTTCT

19407 CGTTACCACAGTCTCCTTGAAACTCAGTCCCCTGACTTGGACTTCCATAACACAATGATT  
TCTGATTTTCTTCTGTTTGTGATTGTTCTTTTGTCCCAGGCACTGGCTACTCCACCTT  
CCACCTCTCTGAAATCATTAGCATTCCCCAAGGATTCTTCAAACTCTCTTTCTTCTTCTG  
GAGAAGTCAGCATAGCTTTAATTTGGACCATTTCTATGGCTTATCTAGATTTTTCAGGA  
CTTGCCTTCAACCTATTCTTTCTGTAGGTGATTCCATTAACCTGTTGCCCATATGGTAGTC  
[C,T]  
GAAGACAGACCTCCGAGAAATGACCCTTGTCTCAAAACTTCCGCAATATGTCCAAATTT  
CCTAGCCTGACATTCAGACTTTGATTATCTGCCTCCAAGTTTATATCCTATCATATTCCT  
TTATATATTCTGTTCTCCAGGTAACTGGGAAGCTTGCCATTCTGATCATAGCCTACAA  
ACTCTTCTGCCTCCCACTCACCTCATCTCTGCTGTCAAAATGCAACCTTCCCTCAAGA

FIG. 3BB

GTCATTTACAGGACCCCTCTTTCTATGAAGCCCTCAGGTGGAAATAATTTTTTGCCTTT

19531 CTCTCTGAAATCATTAGCATTCCCCAAGGATTCTTCAAACTCTCTTTCTTCCTTGGAGA  
AGTCAGCATAGCTTTAATTTGGACCATTTCTATGGCTTATCTAGATTTTTTCAGGACTTG  
CCTTCAACCTATTCTTTCTGTAGGTGATTCCATTAAGTGTGCCCATATGGTAGTCCGAA  
GACAGACCTCCGAGAAATGACCCCTGTCTCCAAACTTCCGCAATATGTCCAAATTTCT  
AGCCTGACATTCAGACTTTGATTATCTGCCTCCAAGTTTATATCCTATCATATTCCTTTA  
[T,C]  
ATATTCTGTTCTCCAGGTACACTGGGAAGCTTGCCATTCTGATCATAGCCTACAACTC  
TTCCTGCCTCCCACTCACCTCATCTCTGCTGTCAAATGCAACCTTCCCTCAAGAGTCA  
TTTCACAGGACCCCTCTTTCTATGAAGCCCTCAGGTGGAAATAATTTTTTGCCTTTTTTT  
CCATTTTATTTTTGGAGTGTTTATGGCATTTAACATACCTTACTTTGTATACAAATATTT  
GCCTTGCTCCCTCTTTTGCAAATTTCTTAAAGGTAGAGACCATTGTATGTTTTCTTCATA

19911 CTCATCTCTGCTGTCAAATGCAACCTTCCCTCAAGAGTCATTTACAGGACCCCTCTTT  
CTATGAAGCCCTCAGGTGGAAATAATTTTTTGCCTTTTTTTCCATTTTATTTTTGGAGTG  
TTTATGGCATTTAACATACCTTACTTTGTATACAAATATTTGCCTTGCTCCCTCTTTTGC  
AAATTTCTTAAAGGTAGAGACCATTGTATGTTTTCTTCATATGTTGCTGGTGCCTAACAG  
AACTATGGCCATTGTCCACATTCATTTAGCAGCCTTTGTAGTTATTGCTTTGAGGAGCTT  
[C,T]  
CTCTCATGAATGCCCTTGCTTTCTCTCCACAGAGTCATCCCCCTATATATGACCTGACT  
GCCATGAAAGTGCCTACTGCTATTTGGGCTGGTGGACATGATGTCCTCGTAACACCCAG  
GATGTGGCCAGGATACTCCCTCAAATCAAGAGTCTTCATTACTTTAAGCTATTGCCAGAT  
TGGAACCACTTTGATTTTGTCTGGGGCTCGATGCCCCCTCAACGGATGTACAGTGAAATC  
ATAGCTTTAATGAAGGCATATTCCTAAATGCAATGCATTTACTTTTCAATTAAGTTGC

20199 TTTGAGGAGCTTCCTCTCATGAATGCCCTTGCTTTCTCTCCACAGAGTCATCCCCCTAT  
ATATGACCTGACTGCCATGAAAGTGCCTACTGCTATTTGGGCTGGTGGACATGATGTCCT  
CGTAACACCCCAGGATGTGGCCAGGATACTCCCTCAAATCAAGAGTCTTCATTACTTTAA  
GCTATTGCCAGATTGGAACCACTTTGATTTTGTCTGGGGCTCGATGCCCCCTCAACGGAT  
GTACAGTGAAATCATAGCTTTAATGAAGGCATATTCCTAAATGCAATGCATTTACTTTTC  
[A,G]  
ATTAAAAGTTGCTTCCAAGCCCATAAGGGACTTTAGAAAAATGGTAACCAACAATGAGG  
TTGTCCCCCAGCACCTGGGGGAGATGCACAGTGGAGTCTGTTTTCCAAGTCAATTGTGT  
TAGTGTTATTTATGTTTAGAGACATCTTTCATGGGACCATCTACAGGTCCTTATAAACA  
ATGAGGTAGATTAGGCAAAAAGATAAACAAGTTGCTACTCTATCTGGCATTTAAGTCTAA  
TTAAATTGTAATTTTAGGGCATACCATGAAGTATAGAAATGTCTGAAGCTTCAAAGGAA

20243 AGAGTCATCCCCCTATATATGACCTGACTGCCATGAAAGTGCCTACTGCTATTTGGGCTG  
GTGGACATGATGTCCTCGTAACACCCCAGGATGTGGCCAGGATACTCCCTCAAATCAAGA  
GTCTTCATTACTTTAAGCTATTGCCAGATTGGAACCACTTTGATTTTGTCTGGGGCTCG  
ATGCCCCCTCAACGGATGTACAGTGAAATCATAGCTTTAATGAAGGCATATTCCTAAATGC  
AATGCATTTACTTTTCAATTAAGTTGCTTCCAAGCCCATAAGGGACTTTAGAAAAAT  
[G,A]  
GTAACCAACAATGAGGTTGTCCCCCAGCACCTGGGGGAGATGCACAGTGGAGTCTGTTT  
TCCAAGTCAATTGTGTTAGTGTTATTTATGTTTAGAGACATCTTTCATGGGACCATCTA

FIG. 3CC

CAGGTCCTTATAAACAATGAGGTAGATTAGGCAAAAAGATAAACAAGTTGCTACTCTATC  
TGGCATTAAAGTCTAATTAAATTGTAATTTTAGGGCATACCATGAAGTATAGAAATGTC  
TGAAGCTTCAAAGGAACAGTGAAATTCCTTTAAGGTCTATATGGAAACCTCTGTTGTCA

20640 GACATCTTTGCATGGGACCATCTACAGGTCCTTATAAACAATGAGGTAGATTAGGCAAAA  
AGATAAACAAGTTGCTACTCTATCTGGCATTAAAGTCTAATTAAATTGTAATTTTAGGG  
CATACCATGAAGTATAGAAATGTCTGAAGCTTCAAAGGAACAGTGAAATTCCTTTAAGGT  
CCTATATGGAAACCTCTGTTGTCAATTTATTTATATGGATTGCTATGGCAATGGACAGAG  
TGTGGGATTAGGAGGAGGGCCTGTAACCTCTTTATAAAAGTTTCTTAGCTATCCTGAAGA  
[T,C]

GTATAGACATTTTTACTTTTTTAGGTATTTTCAACATCAGAAATTCAAAAAAGTCCCCAA  
AGATTCTTCCAGAGAAGCCCTCTTTTCTTACAATCTTATCCCTGGCTATCTGCGTAAACG  
GAATCTTGAACCCATAATAGGATACATGTATAAAATCTTCCTTATTAAAGCAGAAATAAA  
TTGTACAGCATCAATATCATTTTATAATCATAGGGAGGCTTCTTTGTTTAGCATGTAATG  
CCCCCTTACAGGCTTTTGTCTTTGAGGGGTTTGAACATTCCATGAAAACTGACAGA

21156 AGGCTTCTTTGTTTAGCATGTAATGCCCCCTTTACAGGCTTTTGTCTTTGAGGGGTTT  
GAACATTCCATGAAAACTGACAGATAGGAACTGACAATAAAAGATTGAGCTAAAGATG  
GAAGCAGAAAGTACTAGGCTAGATAGTCTCTAAACATTAAGTATTTCTTCTCCATCTT  
AAAAGCAATGAGAAGCCACCAAAATATTTTACCTAATGGAAACCTGATTGCCGCATTTT  
GTAACCACCACTTTGGCTGCTACATAGAGAATGGATTAGAAGATGCCAACAAAAGATTCT  
[G,C]

AGCAAGTCTGTAAATCTGATCAAGTGTTCTGATGCAGGCTGATATCCTTCTGTGCTAAGA  
GAGATGATCCTTGAAAAATCCAGAGCCAGCTCCATAATACTTCTGCTCTGCTGGCAAA  
TCCACAAGCTGCTGGCCCTGGAGCCATTCTTCTCTCAAACTAGCATTCAATTTAA  
TGTATACGTATTGATGGGGAATAATGGTCACTATGAAAACCATGTGATAATATGGAAAAA  
TACCATGATATAATGTTATGTGAAGAGAAGAAAATGAACTGGTAGAACTATGTGATTG

21163 TTTGTTTAGCATGTAATGCCCCCTTTACAGGCTTTTGTCTTTGAGGGGTTTGAACATT  
CCATGAAAACTGACAGATAGGAACTGACAATAAAAGATTGAGCTAAAGATGGAAGCAG  
AAAGTACTAGGCTAGATAGTCTCTAAACATTAAGTATTTCTTCTCCATCTTAAAGCA  
ATGAGAAGCCACCAAAATATTTTACCTAATGGAAACCTGATTGCCGCATTTTGTAAACCA  
CCACTTTGGCTGCTACATAGAGAATGGATTAGAAGATGCCAACAAAAGATTCTGAGCAAG  
[A,T]

CTGTAAATCTGATCAAGTGTTCTGATGCAGGCTGATATCCTTCTGTGCTAAGAGAGATGA  
TCCTTGAAAAATCCAGAGCCAGCTCCATAATACTTCTGCTCTGCTGGCAAATCCACAA  
GCTGCTGGCCCTGGAGCCATTCTTCTCTCAAACTAGCATTCAATTTAATGTATAC  
GTATTGATGGGGAATAATGGTCACTATGAAAACCATGTGATAATATGGAAAAATACCCAT  
GATATAATGTTATGTGAAGAGAAGAAAATGAACTGGTAGAACTATGTGATTGCAATAT

21425 AATGGATTAGAAGATGCCAACAAAAGATTCTGAGCAAGTCTGTAAATCTGATCAAGTGTT  
CTGATGCAGGCTGATATCCTTCTGTGCTAAGAGAGATGATCCTTGAAAAATCCAGAGCCA  
GCTCCATAATACTTCTGCTCTGCTGGCAAATCCACAAGCTGCTGGCCCTGGAGCCAT  
TCTTCTCTCAAACTAGCATTCAATTTAATGTATACGTATTGATGGGGAATAATGGT  
CACTATGAAAACCATGTGATAATATGGAAAAATACCCATGATATAATGTTATGTGAAGAG  
[G,A]

FIG. 3DD

Docket No.: CL001186DIV-II  
Serial No.: (to be assigned)  
Inventors: Gennady V. MERKULOV et al.  
Title: ISOLATED HUMAN LIPASE PROTEINS...

AGAAAATGAACTGGTAGAACTATGTGATTGCAAATATATACAAATATTTAAACAATTAT  
ATGACTTTATAAAATATTTGTATATAATGAAACTGAAGCAATATAAAAAATAAAATTAG  
TTGTGTCAGGGTAGTAACATGATGAGTGATTAATAGTTTTTAATTTTAATATAGTAATG  
ACATAATGTTACAACCTGTCCAAATCTCACAAACATAATATTCAGTAAAGGAAGATAAAC  
ATAAAAGAATACATATTTTATTATACATTTTATGTAGGCTAATTGATGGTTCTGAAAGC

Chromosome map:  
Chromosome 10

FIG. 3EE